PERSONNEL SELECTION BY
STANDARD JOB TESTS

This book is produced in full compliance with the government's regulations for conserving paper and other essential materials.

PERSONNEL SELECTION BY STANDARD JOB TESTS

BY

CHARLES A. DRAKE

Formerly Head, Methods Department
United Merchants and Manufacturers
Management Corporation

FIRST EDITION
FOURTH IMPRESSION

McGRAW-HILL BOOK COMPANY, INC.

NEW YORK AND LONDON

1942

PERSONNEL SELECTION BY
STANDARD JOB TESTS

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

This book offers an exposition of a new technique for the selection of industrial personnel. The method described has been in process of development since 1937. It consists primarily in designing special performance tests that embody the essential elements of dexterity and perception discovered by the analysis of an industrial job or the analyses of a group of similar jobs. Inasmuch as the relation between the tests and the jobs for which they are designed is easily perceived, the method has the advantage of appealing strongly to common sense.

The procedures described imply that the individual who undertakes to use this technique must have a working knowledge of time and motion study, of machine design, of industrial psychology of the conventional sort, of human anatomy and physiology, and of elementary statistics. Any competent employment or personnel worker possessed of adequate training and ingenuity is therefore well equipped to administer such a program for the selection of employees.

The method is still neither complete nor all-embracing. Nevertheless, it is contended that at this stage in its formulation it offers a more realistic and practical method for the selection of personnel than anything heretofore available. It is further believed that greater progress in the continued development of this technique will be made if it is adopted as part of the standard curriculum for the training of industrial engineers than if it is left to persons whose training is primarily psychological and whose interests are largely clinical.

The author wishes to express his indebtedness to the many industrial organizations assisting in the developments set forth. Gratitude is due the Eagle Pencil Company, its officers and executives, for the combination of circumstances that emphasized the initial need for a better technique of personnel selection; Johnson & Johnson, particularly to R. W. Johnson, chairman of the board, for the opportunity to develop and utilize the procedures in their several new factories, in both New Brunswick and Chi-

cago; United Merchants and Manufacturers Management Corporation, especially to Merwin R. Haskel, vice-president, for the further opportunity to employ the method in the staffing of several new textile mills. No less appreciation is due the thousands of employees and applicants for jobs in the nine factories where the tests have been used. Their cooperation and interest were rarely withheld.

The author desires to record the appreciation due other persons who have assisted him directly in highly significant ways. Thanks are extended to Allan H. Mogensen, industrial engineer and nationally known specialist in Work Simplification, for emphasizing the need and practicability of a common-sense approach to the measurement of human abilities; to Clifton H. Cox, industrial engineer, formerly of Johnson & Johnson, for helpful advice and assistance in the practical design and application of the tests; to Betty Martin, formerly of the same organization, for sustained and enthusiastic assistance in the application of the tests; to Holger D. Oleen, industrial engineer, a colleague in the work at both the Eagle Pencil Company and the United Merchants and Manufacturers Management Corporation, whose inventive ingenuity and keen mathematical insight have resulted in valuable contributions to the technique; and to Frances Spodick, formerly executive assistant in the Methods Department, United Merchants and Manufacturers Management Corporation, who has shared in these developments from their first inception and has assumed the major responsibility for their exposition in this form.

The presentation of any new development in scientific management may not justly conclude without acknowledging a continuing debt to Frederick W. Taylor. Every real contribution to this movement is colored by the spirit and zeal which characterized his pioneering efforts. Every worker in this province must admit him as a partner in the work.

CHARLES A. DRAKE.

NEW YORK, N. Y.,
August, 1942.

CONTENTS

PAGE

PREFACE v

INTRODUCTION

MANAGEMENT AS A PRACTICAL ART 1

Status of scientific management. Attitude toward time and motion study. Rating scales and verbal tests. Job tests. Concept of probable error. Relating basic abilities to the job. Plan of presentation.

CHAPTER I

BACKGROUNDS OF JOB TESTING 11

Use of the coefficient of correlation. Early experimentation. The criterion. A new criterion. Selection as science and art. Areas of individual differences. Principles of test design. Training for test design.

CHAPTER II

HUMAN ABILITIES AND INDUSTRY'S PROBLEMS 20

The superior and the average. The inferior. Area of testing. Human structural requirements and machine design. Dual operation. Hand-foot coordination. Rhythm and speed. Inspection and perception. New areas of ability.

CHAPTER III

HAND-FOOT COORDINATION 27

Test results. Another hand-foot test. The machine as a test. Design of fixtures. Other implications.

CHAPTER IV

INSPECTION 35

The illusion of perfect vision. Kinds of visual perceptual ability. Tests for discrimination of small differences. The critical score. Test for span of perception. Intelligence and inspection ability. Illumination and background. Periodic checkups. Testing of other perceptual areas. Inspection and quality control.

CHAPTER V

PAGE

PERCEPTION 48

Early experimental work. Applications of perceptual facts. Social intelligence—an aspect of perception. Salesmanship and perception. Engineering aptitude and perception. Individual differences in perception. Other factors influencing perception. Perception and quality control. Two perception tests. A complex perceptual area. Implications of perception.

CHAPTER VI

DUAL COORDINATION 60

Dual operation ability—a specialized function. Tests of dual coordination. Mechanical aids in dual operation. Savings by dual operation.

CHAPTER VII

A MACHINE-OPERATOR TEST 67

Operation of the test. Test results.

CHAPTER VIII

SOME STANDARD JOB TESTS 75

Sewing-machine Test. Motor Rhythm Test. Hand-foot Test. Hand Dexterity Tests. Inspection Test. Dual Hand Test. Bilateral Hand Test. Conclusions.

CHAPTER IX

MOTOR SKILLS AND PRODUCTION 85

Standard scores. Percentage efficiency scoring. Effects of length of cycles. Increasing efficiency. Monotony. Repetition and fatigue. Ability and the shorter day. Reassigning employees. Selection of foremen.

CHAPTER X

TESTS AND LEARNING ABILITY 94

Sustained production. Effects of low aptitude. The average as ideal. Other handicaps. Fear and accidents. Inhibitions to skills. Effects of brain structure.

CHAPTER XI

APTITUDE AND SKILL 102

Apparatus used. The play. Hypotheses. The procedure. Analysis of results.

CHAPTER XII

PAGE

Measurement of Accident-proneness 113

Early findings. Relationships between tests and accidents. A hypothesis. Implications.

CHAPTER XIII

How to Construct Tests 123

A simple inspection test. Mental operations. Reliability and validity. Critical score. A simple dual test. Special considerations in design. Interpretation of results. Choosing jobs for test design.

CHAPTER XIV

Summary and Forecast 133

Index 139

PERSONNEL SELECTION BY STANDARD JOB TESTS

INTRODUCTION

MANAGEMENT AS A PRACTICAL ART

One might as well concede at the outset of any discussion of management as a practical art that industry exists for the purpose of making profits, and not primarily for the sake of giving jobs, achieving social justice, or even producing goods. It must be further conceded that the production of profits is the primary function of management and that, in the long run, success in the achievement of profits will be the measure against which management is judged.

Management has made rapid strides in the improvement of machines and processes and in the development of instruments for their control. Far less success has attended its efforts in dealing with the human beings working within the industrial structure. Only recently has the human factor been recognized as measurable and predictable within reasonable limits. All science involves measurement, and the weakness of management when it deals with the human factor lies in the inadequacy of its techniques of measurement. However, steps are now being taken to improve and multiply the instruments of such measurement, and this in turn implies progress in the management of men.

Status of Scientific Management

During the early part of this century, the introduction of scientific management by Frederick W. Taylor and others who followed his lead attracted a large amount of attention. Some of the results reported by these early experimenters were astounding. The work of the Gilbreths, particularly, emphasized

the decrease in fatigue that might result from the elimination of unnecessary motions and the use of cycles of operation arrived at by experimental procedures.

Astonishing as these early results were, they failed to stimulate a strong national movement in scientific management. The fine prospect that scientific management might become an art with professional status has not been fulfilled. What remains of the movement is found, in most industries, in the work of time- and motion-study engineers. This is hardly encouraging in view of the fact that industrial engineering has often been boiled down to a single course in the engineering-school curriculum, and time- and motion-study techniques have been taught descriptively rather than embodied in laboratory and field activities.

Attitude toward Time and Motion Study

During and since the First World War, there has been a growing and organized opposition to the use of time-study techniques on the part of certain employers, some labor unions, and many government agencies. The United States Budget for 1940, as presented by the President on Jan. 3, contains this provision on page 742:

> No part of the appropriations made in this Act shall be available for the salary or pay of any officer, manager, superintendent, foreman, or other person having charge of the work of any employee of the United States government while making or causing to be made with a stop-watch, or other time-measuring device, a time study of any job of any such employee between the starting and completion thereof, or of the movements of any such employee while engaged upon such work; etc.

Regardless of how seriously we may view this restriction, which is obviously designed as a safeguard for the millions of government employees covered by the act, it points definitely to two significant conclusions: (1) Time-study engineers have failed to "sell" their techniques in certain high quarters, and (2) there is still sufficient doubt about the scientific validity of their techniques to make safeguards against them a matter of public policy. The redundant effect of such a provision in law

is to bring these techniques under still greater suspicion in private industry.

For this result the time-study engineers are not wholly to blame. Much of their work has been concerned with an area in which measurement was lacking, that of the levels of inherent abilities in the persons whose performances they were measuring. This left their conclusions resting partly on a basis of sound measurement, partly on a basis of assumption and trial and error.

Any adequate solution of their problem involves the conversion of this area of trial and error to one of sound scientific procedures guided by principles evolved from experimentation. The responsibility of the men now practicing as time-study engineers is to inform themselves on what has been done in the scientific aspects of their art, how it has been done, and how they may extend and apply the findings in their own work. The responsibility of engineering schools is to further the basic research, which is now well started and clearly defined, and to make suitable instruction available to their graduates who go into the time-study field.

The trial-and-error area to which particular reference is made is that of human individual differences in abilities and aptitudes significant for industry. As long as jobs are set up with a nice regard for the technical requirements of production but with workers assigned to those jobs on a trial-and-error basis, the work of the time-study engineer will be only partly scientific. It can merit public confidence and scientific approval only when this practice is ended. How may this be done?

It may afford a comforting alibi to assert that the engineer must wait for the psychologist to present the findings of basic research. This is neither true nor helpful. The psychologist has no monopoly on, or vested interest in, such basic research. Since psychological research is largely occupied with other problems, we may wait a long time for such results.[1]

What is required is a better method for selecting the right man for the right job. Only when such a method is in operation can the time-study engineer proceed with any assurance

[1] Adapted from The Time-study Engineer and Aptitude Testing, *Journal of Engineering Education,* vol. 29, no. 10, pp. 856-857, June, 1939, by the author.

that the individuals he is measuring are the ones who are best fitted for the job in question. Moreover, he should know with certainty whether such individuals are merely of average ability for the job or whether they diverge from average by some measured amount.

The numerous excellent texts on time-study techniques are evidence of the progress that has been made in the mechanical aspects of the subject. There has not been, however, a parallel development in direct measurement of the human factor. Skill and effort ratings have, therefore, been used to overcome, at least partially, this lack of exact measurement.

Rating Scales and Verbal Tests

Rating scales have frequently been employed where it has seemed impossible to construct tests. Such scales have certain inherent weaknesses arising out of the fact that they depend upon subjective judgments and lack a common unit of measurement as a basis. The results obtained are qualitative rather than quantitative in nature. Properly handled, however, these results are better than unsystematic, offhand judgments. It often happens that the areas that have been subjected to evaluation by rating methods are later found to be more directly measurable through the use of tests. The latter, in general, are better than rating scales, just as rating scales, in turn, are better than less systematic, subjective judgments.

A good example of the foregoing process is found in the history of intelligence testing. The earlier inferences of an individual's intelligence made from observations of his behavior were later organized into a standard procedure in which scales were employed as measuring instruments. Finally, the scales became standardized and grouped in the familiar test form.

Unfortunately, the measurement of intelligence has contributed comparatively little toward the solution of industrial problems in the way that many enthusiasts had hoped. Those who have employed intelligence tests for the selection of industrial employees have been disappointed at the lack of relationship found between the test results and performances on jobs. Even in the selection of inspectors such tests have not proved useful.

In the selection of management personnel, the use of such

tests has been more promising. The management function requires a high degree of ability in dealing with abstractions. It is concerned with perception of cause-and-effect relations, reasoning, and imagination, all areas in which the conventional intelligence test is discriminative. Sound management practice involves the ability to apply principles in actual practice. While it does not follow that the ability to grasp a principle and express its application verbally is synonymous with the ability to apply the principle in practice, it is nevertheless true that this ability, when identified by a verbal test, may be fairly inferred to foreshadow its carry-over into practice.

The foregoing type of inference has also been found valid in the use of trade-knowledge tests for employment purposes. In these tests an applicant is asked specific questions about the tools, the conventions, and the nomenclature of the trade in which he professes knowledge and skill. The verbal answers are then scored to afford a basis from which an inference is made about his competence as a craftsman. Such tests do not afford a sound basis for a forecast of the individual's capacity for further development. They offer simply an economical procedure for the determination of the applicant's present status.

Certain more generalized tests of mechanical aptitude have been widely employed as a basis for forecasting certain mechanical abilities and for the selection of applicants for instruction. The theory behind the construction of such tests is that mechanical interests lead to knowledge of mechanical movements, mechanisms, and principles, and that such knowledge is, in turn, indicative of underlying mechanical ability. General mechanical aptitude tests have not proved to be useful instruments for the art of management in industry. Many industrial jobs seem to require little of this general ability. The average industrial employee requires abilities that are more largely determined by his anatomy and physiology than by generalized mechanical interest. *Consequently, the tests that will measure the abilities required by industrial jobs must isolate and measure specific aspects of the individual's anatomy and physiology, such as motor coordinations and perceptions.*

JOB TESTS

Motor coordinations include finger dexterities, ability to work with simultaneous or disparate movements of the two hands, with two hands and one foot, with one hand and two feet, or with both hands and both feet. *The measurement of such coordinations can be made only with test mechanisms upon which the applicant can actually demonstrate the degree of control possessed.*

Perceptual measures pose a problem requiring a somewhat different solution. *Perception is defined as the first response after sensation.* Thus, when we see an orange we see a colored, spherical object. Our next response is to identify this object as an "orange" rather than as a colored ball, a grapefruit, or an apple. This second response is the perception and gives meaning to the original sensation of sight. The perceptual process rejects the inappropriate meanings and accepts the item most appropriate.

In measuring perception, we are in constant danger of confusing the perceptual response with the first or sensory response. This may lead to incorrect inferences and disappointments in practice. For example, for a long time it has been thought that perfect eyesight is a guarantee of perfect perceptions. Research has shown that perfect vision affords no such guarantee, that with persons of normal vision there is a wide range of ability in the speed and accuracy of perception. This is discussed in detail in Chaps. IV and V.

This distinction between sensation and perception must be observed in the construction of tests and the interpretation of test results. Poor vision is not disqualifying for certain jobs. Nearsightedness may be a positive asset, because under certain conditions it is shown that this condition favors perception of the sort required in close and fine handwork. On the other hand, certain eye defects impose a marked handicap upon the speed and accuracy of perceptions. In the practical work on the job, we are more concerned about the quality of perceptual ability than about the eyes as organs of sensation.

Concept of Probable Error

Since no measuring instruments are perfect, the measurements made are always accurate within determinable limits. These limits are statistical concepts and require interpretation in terms of the theory of probability. In other words, the measurements made must be interpreted in terms of their probable error. *This probable error expresses the number of chances out of* 100 *that the true value sought by the measurement lies between certain limits above and below the measurement obtained by the use of the instrument.* Thus, if a measured value is found to be 112 points, and the probable error is expressed as 8 points, the chances are 50 out of 100 that the true value lies between 104 and 120. Within this range, the chances are greatest that the value will lie at, or near, 112. Chances are much less than 50 out of 100 that the true value will lie below 104 or above 120, although there is still a probability that can be expressed in statistical terms that the true value might be as great as 144 or as small as 80. This concept of probable error, due both to the variability of behavior or performance as observed in the same individual from one point in time to another point in time and to the differences in performances between different individuals, is implicit in all handling of the human factor.

All predictions of human performance are thus shown to be subject to some errors arising out of the nature of the human organism, the imperfections of the measuring instruments, and the inadequacies in the techniques employed. But the amount of such errors is determinable within limits, and the implications must be thoroughly understood. This concept of probable error has long served as the basis for insurance premiums and must be made the basis for the forecasts of productive accomplishment in industry.

The early exponents of scientific management were not oblivious of this fact of probable error. They did lack, however, an adequate technique for the measurement of human abilities which would yield norms around which probable error might be computed. Failure to evolve such techniques may be blamed upon the slow development in the area which psychology had claimed as its special province. There was no particular reason

why the task should have been left to the psychologist. Considering the academic preoccupation of the latter, it would seem rather more reasonable to have had the basic experimental work performed on the job by the practitioners of time and motion study.

Relating Basic Abilities to the Job

There was another factor, rather more prevalent during the first two decades of the century than now, that probably operated to inhibit the necessary experimentation. This was the

Table I.—Pin Board F

Time scores, seconds	Applicants	Operators	Operators ranked by foreman
120–129	37	1	(31)
130–139	81	2	(2) (4)
140–149	79	6	(6) (10) (12) (13) (26) (29)
150–159	51	9	(1) (3) (5) (11) (15) (25) (30) (32) (40)
160–169	54	8	(7) (8) (9) (14) (17) (19) (23) (24)
170–179	39	4	(21) (22) (28) (33)
180–189	17	2	(18) (20)
190–199	9	4	(16) (27) (34) (35)
200–209	1	1	(36)
210–219	3		
220–229	1		
230–239			
240–249	1	2	(37) (38)
250 up		1	(39)
Totals....	373	40	

view that the range of human motor abilities was so flexible that exceptional skill could be achieved merely by a favorable combination of willingness and intensive training. Experimentation within the last few years has largely dispelled this erroneous view. Testing of basic abilities has now arrived at the point where the measurements obtained show a highly satisfactory relationship to success on the job. As an example of one such achievement, examine Table I, above. This gives the results of the application of such a test, shown in Fig. 8,

page 61, to 413 operators. The scores on this test range from 120 seconds to more than 250 seconds, *i.e.*, there were some persons who required more than twice as much time to perform the test as others. If we consider 150 seconds as an acceptable performance, there is a substantial number of persons who can better this record by 20 per cent and another substantial proportion who will require much more time. Whereas this test is extremely significant for identifying the average and above-average performers, its greatest utility, from the standpoint of production, lies in its power to sift out poor operators, as shown in column (4). One may correctly infer from this that the applicants who made scores of 170 seconds and more are potentially poor operators and should not under normal conditions be employed for this type of work.

It should be noted that, of the 14 operators having scores of 170 seconds or more, 11 were in the lower half of the foreman's rankings. One might say that the odds are something like 4 to 1 that the applicants with such scores will become substandard operators if they are employed. Certain confidence may be placed in the reliability of these rankings by the foreman, since they are largely based upon his knowledge of the piece-rate earnings of these operators.

This particular test does not show as great a range of ability as do other tests that have been constructed. On some of the latter the slowest operators required six times as much time as the speediest operators.

The time-study engineer must know the basic ability of the person whose operation he is studying. If he can be sure that this person is at the norm, or 100 per cent level, or that he is at 120 per cent or at 80 per cent of normal, he can establish his standard times with assurance. If, however, the basic ability of the operator is unknown, the time-study engineer can only make his estimates subject to much larger probable errors. If human ability is to be placed for its own best advantage in industry, and management is to use such human ability most effectively in the production process, a continuous development of aptitude tests and their further extension in selection, training, and placement is imperative. This obligation rests directly upon the technical personnel, the engineers, and the industrial psychologists, but the establishment of policies and the provi-

sion of facilities and supervision for such procedure is an inescapable obligation of management.

Plan of Presentation

This book presents a set of techniques for the measurement of the human factor, with constant regard for the fact of probable error in its measurement and probable error as applied to its predictions. The plan is to offer, first, a consideration of the backgrounds of job testing, both of the usual sort and of the type developed by the author. This is followed by a discussion of human abilities as they bear upon problems of industrial selection and training. The next several chapters deal with special areas of ability, such as hand-foot coordination, inspection, perception, and dual operation. Attention is then devoted to certain specialized tests for sewing-machine and other machine operators and for a variety of other jobs involving dexterities. Also considered are the general problems of motor skills, learning ability, accident-proneness, and other measurable characteristics of human nature as they bear upon the problem of test design and construction. A chapter entitled How to Construct Tests gives specific directions for the construction of performance tests in accordance with the principles elaborated in the text. In conclusion, there are a brief summary of the accomplishments thus far and a forecast of what the further development of this method holds forth for industry.

In no sense is this book a historical treatment of the attempts to measure human abilities. It is rather a concise presentation of the principles that may be employed in the design of effective instruments for the measurement of those abilities which are significant in industry. It is a guide for practice rather than a well-rounded exposition of the problems involved. The techniques described are offered as a new instrument in the further progress of scientific management.

CHAPTER I

BACKGROUNDS OF JOB TESTING

For many years, the customary procedure in industrial psychology has been to apply a battery of from 3 to 20 standardized tests, mostly of the pencil-and-paper variety, to a sample of employees on some particular job. These employees were arranged in an order of excellence from "best" to "poorest," usually by foreman's or supervisor's rankings, piece-rate earnings, or by quantity and quality of production, to constitute a criterion against which the efficacy of the tests could be measured.

By complicated statistical processes, the tests having some significant relationship to the criterion were identified and combined into a smaller battery, usually consisting of from two to six tests, which could then be applied to new applicants for purposes of selection.

This procedure was cumbersome and costly. The testing time per employee was excessive. Materials were often expensive, and the services of a trained psychologist who was also a competent statistician were required. The whole process was often viewed with suspicion by management, and particularly by industrial engineers and employment personnel.

Use of the Coefficient of Correlation

The *coefficient of correlation* was generally employed to express the relationships between test scores and the criterion measures. A perfect relationship, in which the successive test scores, from highest to lowest, are exactly paralleled by the successive criterion measures, from highest to lowest, would be expressed as $+1.00$. The inverse situation, in which the measures in descending order in one series are exactly paralleled by the successive scores in ascending order in the other series, would be expressed as -1.00. The complete absence of either

relationship would be expressed as .00. A coefficient of correlation of ±.90, or more, between test scores and criterion would be considered high; between ±.60 and ±.80, moderate; between ±.30 and ±.50, low; and ±.20 or less, negligible. Since the probable error of predictions increases rapidly as the coefficient of correlation diminishes, only the highest correlations are useful in practical situations, and these are seldom obtained.

In fact, in the customary procedure, the size of the coefficient of correlation between the final battery of selected tests and the sample of employees used as a criterion was usually not impressive, often ranging from only .30 to .50. Moreover, some of the tests remaining in this selected battery, when viewed in the light of ordinary common sense, seemed to have no relationship to the nature of the work performed.

Early Experimentation

The writer began his work in industrial psychology in 1936, thoroughly schooled in the foregoing techniques. He had already published a study in a related field, involving more than five years of experimentation. For some years previously he entertained serious misgivings about the method as a feasible approach to the intensely practical requirements of industrial employee selection. These misgivings were heightened by reports of work in which the job required dual hand operation but for which the test battery included not a single test of such two-hand coordination ability. He was astonished to find employees on jobs requiring tactual perception, kinesthetic perception, and auditory perception tested with batteries which included not a single measure in any of these areas.

In the early experimentation at the Eagle Pencil Company in 1936–1937, it soon became apparent that it was quite possible to take the analysis of the operation cycle on any job, as made by time- and motion-study engineers for methods and rate-setting purposes, and use it as the basis for the design of specialized tests for that particular job. It was concluded that the extensive Therblig Analysis, made by the Gilbreths and their successors and now used in all modern industrial engineering, particularly in the Work Simplification training program of Allan H. Mogensen, could be reduced to five basic elements for purposes

of test design. These elements are: (1) Select and grasp; (2) transport loaded; (3) position and release; (4) any intervening element (pressure on a foot pedal, use of an operating lever, etc.); and (5) transport empty.

In addition to the standardized tests which were at first used extensively in the Eagle experimentation and later gradually abandoned, nine special tests were designed for jobs and groups of jobs in the Eagle plant. The testing of hundreds of employees and applicants yielded invaluable basic data upon which the technique of design has since been elaborated. Well-marked areas of human individual differences were identified; and these specialized innate abilities were shown to be closely related to performances on jobs. The primary analysis indicated that most jobs could be classified into the following categories for purposes of test design:

1. General hand dexterity.
2. Dual hand coordination.
3. Hand-foot coordination.
4. General body coordination.
5. Perceptual ability.

Subsequent work showed that these categories could be still further broken down when the number of jobs in the subordinate categories justified the design of more specialized tests.

The Criterion

This early experimental work also showed that it was difficult, if not impossible, to obtain a satisfactory criterion against which to measure the effectiveness of a given test. Each of the possible criteria may be examined in the light of this investigation:

1. *Foreman's Rankings.*—Practically all foremen are subject to personal bias against some of their workers. They also suffer from the "halo" effect, *i.e.*, the tendency to overrate individuals who have shown an exceptionally good or pleasing performance at some time or other. They identify their best workers and poorest workers with some discrimination but are quite incapable of giving a consistent or obviously verifiable ranking of employees around the average in ability. At best this ranking is on a general worth basis rather than on the basis of in-

spection ability, dual operation ability, hand-foot coordination ability, etc., which is required in test construction. When the foreman's rankings are based upon his knowledge of production or earnings, they have higher reliability, without a marked increase in validity.

2. *Production.*—The best workers seldom or never produce at their fullest capacity, for reasons that are generally familiar. The poorest workers are aided by others to bring their production up to some acceptable standard. And the average workers plod along at some stereotyped output, consciously or unconsciously arrived at, which does not endanger either the piece rates or the job.

3. *Earnings.*—These figures generally reflect production and are subject to the same influences. In addition, they are often modified in favor of age and length of service of particular employees.

4. *Percentage Efficiency.*—These figures, arrived at by time-study techniques, are liable to fewer subjective errors than any of the foregoing, but are far from constituting a reliable criterion—a conclusion with which every honest time-study engineer is bound to agree.

5. *Quality.*—Some weighting of quality of production is often combined with one or more of the other criteria. The difficulty here is in getting an objective measure of quality. Subjective appraisals of quality are notoriously unreliable, whereas objective measures, based on rejections, spoilage, and waste, are often modified by causes wholly beyond the control of the operator.

A New Criterion

In view of the foregoing situation with respect to a criterion, the writer ventures to offer a proposal that is an inversion of the usual point of view. It is now established beyond question that a properly constructed test which measures basic innate abilities called for on a particular job or a group of similar jobs gives a highly reliable classification of employees with respect to these abilities. Such a test, applied repeatedly and after different intervals of time, gives relatively the same results, practically unmodified by experience on the job.

It is proposed, therefore, that the test results be used as the criterion against which the efficiency of the productive process may be measured. This amounts to saying: "Here is an employee of average ability; you should get average production from him. Here is an employee of very superior ability; you should get very superior performance from him. Here is an employee much below average in ability; you must make up your mind whether or not you want him on this job."

This places the responsibility for performance squarely upon management, and particularly upon supervision. Employee ability is ascertained beyond question. The problem of realizing that ability in terms of production is strictly a responsibility of management.

Selection as Science and Art

As methods engineering becomes more precise as an art, it supports the contention that management, in general, is not getting the best results from its superior employees, whereas it tolerates and nurses along other employees on unsuitable jobs. Management idealizes average performance at a level of statistical mediocrity, whereas its best chance of surviving in a competitive society comes from making the best possible use of the most superior employees it can obtain through suitable testing procedures. The advance in methods engineering as a science and an art has eliminated this as a responsibility of supervision. It has become clear that the burden is definitely upon methods engineering as a science to furnish precise measures of employee abilities, which may then be utilized in the productive process. Methods engineering stands, therefore, in exactly the same relation to management as do physics, chemistry, and the other sciences.

It follows from this analysis that methods engineering must be reduced to a body of experimentally verifiable principles. Just as other sciences develop upon a basis of verifiable data and measurable interrelationships among phenomena, so does methods engineering rest upon a body of ascertainable facts about human individual differences, and the relationships between the abilities they represent and the results of the productive processes.

Having such a body of verified principles, methods engineering as an art may proceed to apply such principles to the improvement of the productive processes. As such, it becomes one of the essential tools of management, a necessary staff function.

Areas of Individual Differences

It is now possible to state in this stage of advancement of the science that the following areas of individual differences are known, are measurable, and are related to production:

1. General finger dexterity.
2. Bilateral hand coordination.
3. Dual hand coordination.
4. Dual hand-and-one-foot coordination.
5. General body coordination, including ambulatory speed and dexterity.
6. Rhythm, in motor and possibly perceptual areas.
7. Perceptual ability specialized with respect to sensory areas, such as visual, auditory, tactual, etc.

The foregoing list is not complete. New areas will undoubtedly be discovered in further experimentation. It will, however, serve as a guide for the design of tests and for the classification of jobs for which those tests are appropriate.

Principles of Test Design

As for the tests themselves, it is now possible to formulate a body of principles which apply to their design:

1. The test must measure innate individual differences, and not learned skills.
2. The test must include the elements or "therbligs," particularly the positioning element, in the proportion found on the job.
3. A separate test, or tests, must be used for the measurement of perceptual factors.
4. The test must be long enough to give a reliable measurement.
5. The test must be easily comprehended, and a practice period must be allowed to assure complete familiarity with the cycle of operation.

6. The test must be as well designed, *i.e.*, designed in accordance with up-to-date engineering principles, as is the job setup for which it is appropriate.

7. The test must be safe in operation, removing any danger of injury to the testee and preventing any fear on the part of the applicant being tested.

8. Motor tests should preferably be of the work-limit type, with a certain amount of work to be done or pieces to be placed, *i.e.*, a set number of cycles of operation to be completed, the score being taken in terms of time required to complete this amount of work.

9. In perceptual testing, the items of the test must be graduated in order of difficulty from items that anyone can discriminate to items that are so difficult that they are clearly beyond the limits of perception of the very best operators. Such tests must be scored in terms of both time required and accuracy achieved.

10. The test must be economical to construct, with a cost that is not out of proportion to the service it is rendering.

11. The test must be durable, *i.e.*, little affected by the wear and tear of repeated use.

12. The test must be easily repaired, with parts easily replaceable.

13. The test should be flexible, so that it may readily embody any improvements and changes in the job.

14. The test should be capable of application by the usual methods and personnel employees, and not require the services of a trained psychologist.

15. The test must be easily scored in terms that anyone can understand.

16. The test must be compact, not a bulky piece of equipment.

17. The test must be quickly reset, ready for the next testee without laborious rearrangement by the tester.

18. The test must be amenable to common sense, *i.e.*, the relation between what is apparently measured by the test and what is apparently required by the job should be easily perceived.

The design of tests is itself an art, which must undergo a process of development like any other art. It requires certain

skills and certain insight which are already well known in the field of machine design. The individual who undertakes to design industrial aptitude tests will find his task much easier if he is capable of producing his designs in wood and metal, ready for operation. If he must sketch his designs and then merely supervise their realization in the test model, he will lose the advantage that comes from continuous experimentation during the process of construction.

Training for Test Design

It must be obvious from the foregoing that the individual who attempts to design these aptitude tests must train himself in areas beyond the usual courses in academic psychology. He must have some command of statistics, of sketching and drafting, and of woodwork and metalwork, particularly of the operation of power tools, such as lathes, band saws, drill presses, shapers, circular saws, and sanding machines. If he is to turn out a finished product, he must know something of strength of materials and the elements of machine design. He will find it advantageous to know something of photography and to be able to take both still and motion pictures of the job and of his tests.

His personality should be such that he can quickly gain the confidence of the workers on the jobs, and secure and maintain the respect of higher management, the engineers, and the supervisors with whom he must work. He should be both willing to learn and capable of learning to operate a machine or fixture under the tutelage of the regular operator, and to work at a bench on the job he is studying. This firsthand experience will go far to facilitate insight into the problems presented by the job.

Such an individual is neither a super-engineer nor a super-personnel department. His primary duty is that of creating a tool and making it available to the organization that must use it. It is his obligation to make his technique comprehensible to the personnel who must employ it. It is thus quite essential that he comprehend the problems and points of view of management and industrial engineering, that he be alert to changes occurring in these fields, and know personally or by report the leading figures and their work in these areas. He should add

to his vocabulary the terms and nomenclature current in these fields. He must be able to talk the language of the manager, the foreman, and the engineer, employing terms with which they are familiar and for which they have precise meanings. Only in this manner can he justly expect the cooperation necessary for the fullest development of his project.

CHAPTER II

HUMAN ABILITIES AND INDUSTRY'S PROBLEMS

Vocational guidance counselors and others concerned with the guidance of young people into vocational activities often express regret at the lack of opportunities in industry for the utilization of the abilities observed in the persons they attempt to advise. They remark particularly the lack of opportunity for self-expression. It has even been proposed that industry be generally reorganized to yield more opportunities for such self-expression on the part of those who work in it. Whereas such a proposal cannot be taken seriously, it does serve to emphasize an attitude that is well established in many people who deal with educational problems. It serves, too, to bring into prominence a tendency to look upon the means of livelihood as a dominant part of life. The view would have more validity if the time spent in industry amounted to 12 or 14 hours of a 24-hour day.

Against this view one may suggest that the time actually spent in working is only a part of the total time available to the individual; that he may, through avocational activities, through his outside social life, and through his recreations, find ample opportunity for self-expression. This implies that the vocation should be looked upon as a means to an end—that means being self-support—and not an end in itself. To achieve the highest possible effectiveness in the cooperative effort of industry, it is obviously important to afford the workers an opportunity for adequate self-expression; but just as democracy itself can operate only through the surrender of functions by the individual, so industry can operate only through the surrender to management of the functions appropriate for the effective combination of labor with the other instrumentalities of production.

The Superior and the Average

One of the chief laments of the vocational guidance specialist is that many persons of superior intelligence are forced to take jobs requiring far less intellectual ability than they can offer, because the number of jobs requiring a higher order of mental ability is sharply limited. True, intelligence of a superior order is not usually called for in high degree below the ranks of management. However, it may be just as illogical to ask industry to provide an opportunity for the exercise of superior mental ability as it would be to ask industry to provide outlets for filial affection or religious devotion. Life provides suitable opportunities for these elsewhere.

From the viewpoint of industry, it seems necessary to effect an organization around average individuals. While this may lead to a lower order of performance than could be achieved by selection of the best, it does make for permanence of the organization and easy replacement of individuals lost to it. It makes standardization of supervisory activities and training procedures relatively easy. Superior individuals require more skill in handling and consequently make more demands upon the ability of management. This, in turn, implies that management itself must be in the hands of highly superior individuals.

The tragedy of the situation is not that the individuals of superior intelligence are denied employment, but that they are scattered indiscriminately over the jobs, often in situations where their ability cannot be applied to the best advantage. At one time the writer was called upon to organize the training of military personnel to replace trained civilians in army schools. By a systematic canvass, he found a considerable number of college, high-school, and trade-school graduates. Many of them had concealed their level of training for reasons which seemed quite sufficient to them. Such training seemed of no particular significance for a private in the ranks. By careful selection, appropriate training, and suitable encouragement, they formed a wholly competent staff of teachers for other enlisted men.

In general, the possession of superior intelligence would not offer, in itself, a serious problem in industry. There is known

to be a much greater turnover among the more highly intelligent who are placed on routine jobs. If the intelligence level is known, or can be inferred from previous attendance at high schools and colleges, it is often better to reject an applicant for this reason alone.

The Inferior

At the other extreme lie the individuals of markedly inferior abilities, occasionally presenting industry with representatives of the moron and imbecile groups. It is not possible to classify these persons accurately without some test procedure. It has been observed, however, that their failure to comprehend simple instructions, their inability to follow directions, and their marked lack of attention are symptoms to arouse suspicion at the time of employment.

The individual with low-grade mentality is often a danger to his own life and limb and a menace to others. His lack of insight into cause-and-effect relationships renders him oblivious of danger in many situations; his lack of foresight precludes suitable action to avert a serious consequence; and his inattention makes him careless and unreliable.

In the processes of training, the stupid individual shows two marked characteristics. He is slow to learn and often apparently clumsy in his physical coordinations as well. Also, he is exceptionally slow to break established habits in order to form new ones. It follows that such individuals are difficult to train, that wrong methods of work, once established, are very difficult to overcome, and that the individuals require considerable supervision and minute directions. While they may be suitable for simple tasks requiring strength rather than dexterity, great care is necessary to prevent their assignment to jobs beyond their limited abilities.

Area of Testing

When one speaks of testing the basic abilities called for in industry, the usual assumption of the manager or engineer is that intelligence tests are implied. Another assumption is that the test must be printed on paper and performed with the aid of a pencil. It is now well established that tests of the pencil-

and-paper variety have very limited utility in selecting operating personnel for industry. Aside from intelligence and certain aspects of perception, the abilities to be measured call for the use of performance and manipulative tests. It is with tests of this latter character that we are primarily concerned. As was pointed out in the Introduction, information tests on procedures, nomenclatures, and descriptions of operations have a small amount of diagnostic value in clerical occupations and in the trades. These should, however, always be supplemented by a practical test, involving the use of the tools of the craft, for a valid measure of ability and capacity for development.

Aside from intelligence, there is a large number of other individual differences having a direct bearing upon successful performance in industry. Some years ago, a classified advertisement in a newspaper asked for a stenographer not over 5 feet tall and weighing no more than 90 pounds. Upon investigation, it was found that the job involved working in a small space under a staircase, and that a person of normal size could not work comfortably under this condition. In another instance, a company employed stockroom personnel without regard to size. It was found that several very short employees required more than twice as much time to assemble certain articles that were stored on shelves beyond their reach, because they had to bring a chair or a stepladder to the point of storage.

Both the foregoing cases emphasize the gross physical requirements that jobs sometimes impose. Many situations exist in which the tall person is handicapped and many others in which the short person is at a disadvantage. There are jobs, such as the operation of a standard typewriter or sewing machine, where the setup favors a left-handed person; jobs favoring persons who are deaf, blind, or mute; and other jobs totally inappropriate for individuals impaired in any of these latter functions.

Human Structural Requirements and Machine Design

The writer has seen machines so designed that the operators worked cross-handed. This, of course, was poor designing—the result of preoccupation with mechanical movements without due attention to the limitations of the human structure. He

has seen any number of jobs set up with fixtures arranged in such a way that the pickup and positioning elements were imposed upon the left hand. This would be quite appropriate with left-handed operators, but it imposes a hardship entailing sacrifice in speed upon right-handed operators.

Dual Operation

It is now well established that wide differences exist in the ability to coordinate the two hands. While dual operation should yield an increase in productive efficiency of around 30 per cent, this can be expected in not more than one out of five persons. There is still some question whether one or more of the other four persons can be trained to effective dual operation. The writer's experience points to the conclusion that, if such training is possible, it can only be achieved at a cost of time and effort usually not justified by the result. The simpler procedure is to select by tests the individuals with the basic dual operation abilities.

Hand-foot Coordination

In the coordination of hands and feet, as on foot presses and foot-operated fixtures, there is known to be a wide range of ability. It is often highly hazardous to permit the assignment of individuals lacking this ability to machines that require it. This suggestion will be discussed at length in a later chapter, with illustrations from experiments.

Rhythm and Speed

Still another area of wide individual differences is that of rhythm of work. Whereas it does not follow that the possession of the ability to feed a belt-feed or a dial-feed machine rhythmically and with precision is related to ability in music, it has been shown that such ability can be satisfactorily measured by appropriate performance tests.

There is danger that the ability to work rhythmically may be confused with the optimum speed of performance. *Optimum speed of performance is the most comfortable speed at which the*

individual can continue to operate. This is easily illustrated by a commonplace experience. If you are walking or driving behind an individual whose pace is too slow for you, you become annoyed and impatient. On the other hand, if you are attempting to keep up with an individual who is walking or driving at a pace you find difficult to maintain, you are similarly under a strain and soon become exhausted. The parallel case for rhythm is illustrated by walking or driving behind a person who slows down, speeds up, and slows down again with no regularity. Lack of rhythm may be fully as annoying and exhausting as the strain of a pace that is too fast or too slow. The sound solution in either case is to find an individual whose sense of rhythm is appropriate to the work and whose optimum speed is the standard pace of the work.

Inspection and Perception

Inspection ability rests upon basic individual differences in perception. There is a perceptual ability for each of the senses, but, as has been pointed out, this ability is not a direct function of the efficiency of the sense organ involved. Perception is a central or brain function and, so far as is known through experimentation, is not significantly modified by experience or training. It is important, therefore, in selecting individuals for inspection work to choose for training only those who possess the basic perceptual abilities called for by the work.

New Areas of Ability

Finally, it must be pointed out that there certainly must be a large number of individual differences still undiscovered and unmeasured, which will emerge from time to time as experimentation continues. Many of these will be extremely significant for skilled and semiskilled operations in industry. Delicacy of coordination, sequences of skilled movements, and combinations of movements with perceptual acts must be considered both in the design and setup of machines and jobs, and in the testing of the individuals to be selected as operators. When actual industrial performances show some individuals able to

produce comfortably at a rate 200 per cent higher than some other individuals, it is high time to develop techniques to identify these far superior individuals and to secure their services while our plants operate competitively in a society dominated by the profit motive.

CHAPTER III

HAND-FOOT COORDINATION

It has long been tacitly assumed that general hand dexterity was closely associated with other bodily dexterities, and that a test of such general hand dexterity would be quite adequate for the selection of employees for jobs involving one or more foot elements in the operation cycle. It is now known that this is not the case. Hand-foot coordination within a single operation cycle shows a wide range of individual differences.

Test Results

The figures from the Eagle Pencil Company experimentation, shown in Table II, give the intercorrelations of the several tests used in the study. It may be noted that the intercorrelations of the hand-foot test scores with the largely manual test scores have

Table II.—Intercorrelations of Aptitude Test Scores

Test	PB	90°	L-R	R-R	H-F	MR	CT	CE	ST
90°	.59								
L-R	.48	.26							
R-R	.43	.26	.59						
H-F	.58	.41	.40	.41					
MR	−.22	−.11	−.11	−.13	−.28				
CT	.36	.33	.30	.16	.38	.10			
CE	−.03	−.01	−.12	.04	.05	.01	−.22		
ST	.43	.27	.28	.14	.40	−.08	.39	.02	
SE	−.05	.07	.03	.03	.07	.04	.10	−.05	.03

Code

PB.......... Pin Board
90°......... 90-degree Dual
L-R......... Left-right Dual
R-R......... Right-right Dual
H-F......... Hand-foot Coordination
MR......... Motor Rhythm
CT.......... Case Inspection Time
CE.......... Case Inspection Error
ST.......... Spiral Inspection Time
SE.......... Spiral Inspection Error

a range from .40 to .58. The hand-foot test used in this experimentation is shown in Fig. 1, below. On this test the hand operation was bilateral. The right hand picked up a large peg and inserted it in the crossbar, while the left hand picked up a small peg, inserted it in another hole at right angles to the larger hole, and pinned the large peg fast by insertion of the small peg through an eye in the large peg. The foot lever was

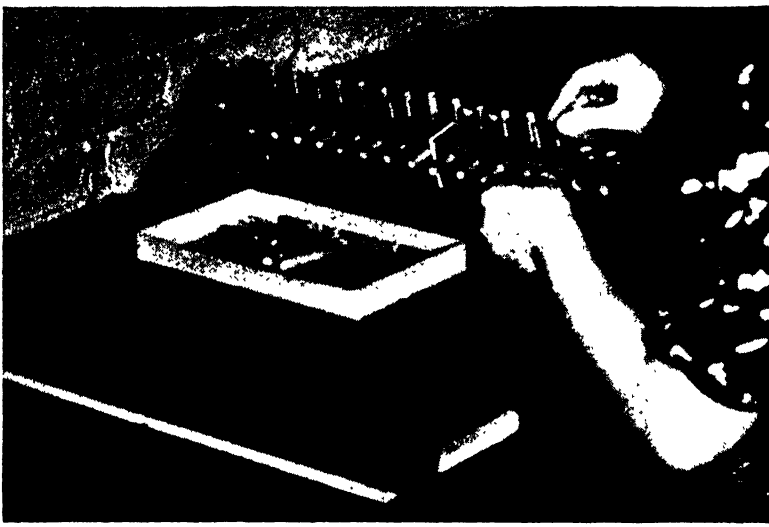

FIG. 1.—Hand-foot Test I.

then depressed, giving the bar a one-quarter turn, and the hand operations were repeated. The foot lever was then released and the bar returned to the original position for the next set of hand operations. The range of scores on this hand-foot test is shown in Table III, page 29.

Experience in attempts to train persons with poor scores on this test on foot-operated fixtures was discouraging, the final level of proficiency attained not justifying the training effort. Retests of employees on this test after an interval of several months gave substantially the same scores as those attained on the first test.

Another Hand-foot Test

In the Johnson and Johnson experimentation, the hand-foot coordination test involved dual hand operation, the two hands performing identical elements simultaneously. On this test, three exposed pairs of holes were filled from top to bottom with small round pegs arranged parallel to the near edge of the test, flat wooden pieces at right angles to the near edge of the table, and half-inch steel balls in round holes. The foot lever was

Table III.—Distribution of Scores on Hand-foot Test I

Time, seconds *	Number of persons	Percentage efficiency
95–109	7	143
110–124	24	125
125–139	37	111
140–154	20	99
155–169	12	90
170–184	12	83
185–199	8	76
200–214	4	71
215–229	2	66
230–244	1	62
245–259	1	58
	128	

* Average = 146.68 seconds

then depressed and held down, while two more steel balls were dropped in two newly exposed round holes. This test is shown in Figs. 2*a* and 2*b*, pages 30 and 31.

The range of the intercorrelations of scores on this test with the other manual dexterity tests used is from .40 to .62, and thus agrees quite closely with the results of the Eagle hand-foot test.

Neither of these tests would be ideal for the selection of operators for the usual foot presses and foot-operated fixtures. The Eagle test was designed for jobs requiring the positioning of the parts manipulated partly by sight and partly by touch in a direction at an angle of 90 degrees to the median plane of the body. The Johnson and Johnson test was designed for a dual

winding operation in which a number of these positioning elements occurred in the operation cycle.

Most foot-press jobs involve positioning by only one hand, followed by depression of the foot lever and its immediate return to the upward position, while one or both hands remove the

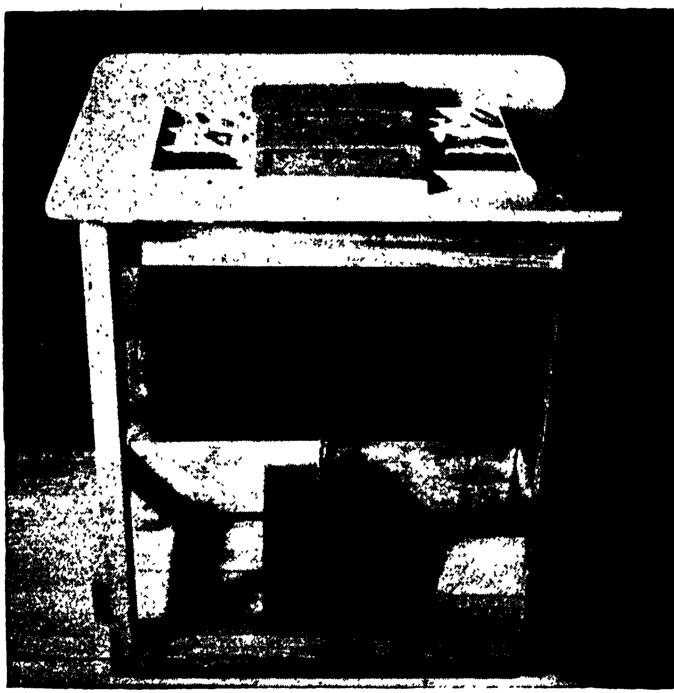

FIG. 2a.—Dual Hand-and-foot Test.

work from the die. Many of these operations on the standard foot presses and kick presses could be more economically performed by dual operation. Where such dual operation is called for, an appropriate test will involve analogous dual hand operation and a foot element inserted at the proper place in the operation cycle.

THE MACHINE AS A TEST

The question has sometimes been raised, "Why not test applicants on the machine itself, rather than on a specially de-

signed aptitude test?" The answer is more obvious where the foot press is involved than where the job involves only the use of hand tools or simple fixtures. The primary objection is that the press is a production instrument, designed for a special function, consuming more or less valuable materials in operation, and more or less dangerous. An especially designed test emphasizes only the important elements in the operation cycle

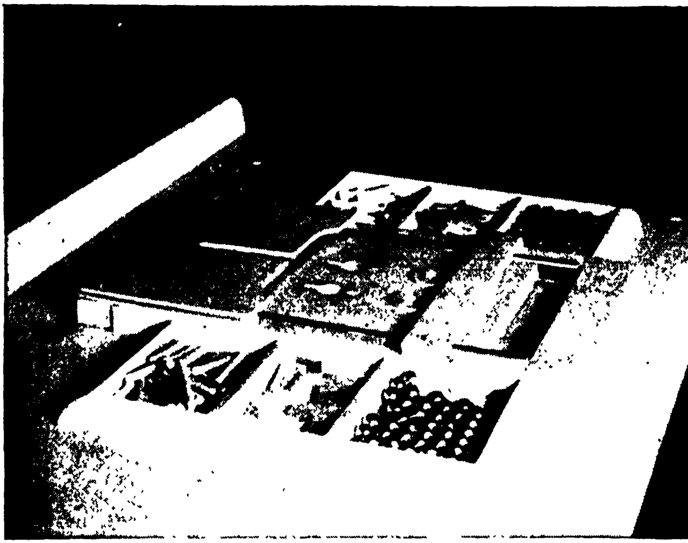

FIG. 2b.—Dual Hand-and-foot Test.

largely independent of acquired skills, uses no production materials, and is, or should be, perfectly safe in operation. This last factor is of considerable importance, since accidental injury to a person not on the pay roll of a company might have serious consequences. It is this last factor that makes it highly inadvisable to use power machinery for the discovery of aptitudes. It should be clear that a hand-foot coordination test that is to meet adequately the needs of a plant using many foot presses and many foot-operated fixtures must be quite flexible and, at the same time, quite simple. Its operation cycle should contain positioning elements as extensive as the positioning elements commonly found in the job, and the foot element should occur in the corresponding place in the operation cycle.

DESIGN OF FIXTURES

Among some industrial engineers there has been strong objection to job setups involving two foot pedals, the view being that one foot resting on the floor was necessary to steady the operator and to prevent undue fatigue. This view is largely obsolete. Certain types of power sewing machines as well as certain specially designed fixtures involve the use of two similar foot pedals. These have contributed markedly to the more effectual operation of the mechanisms. On most of the present job setups involving hand-foot coordination, additional hand elements are introduced to remove the work from the press or fixture. In most instances, such removal could be more quickly accomplished by a drop-delivery mechanism actuated by the other foot, thus freeing the hands to begin the next cycle.

It is clear, of course, that, if the job is set up for bilateral foot operation, the test must involve provisions for testing the appropriate coordinations. As the complexity of the operation cycle increases, the complexity of the test will increase, and accordingly the proportion of applicants who will meet satisfactorily the coordination requirements of the job will be considerably smaller.

It is not unusual to find a job setup with complex, special fixtures involving a complicated series of hand-foot operations, which, though quite ingenious from the standpoint of methods engineering, may have lost many of the advantages that are possible through suitable division of labor. This is especially likely to be true when the selection of operators is made without the aid of appropriate tests.

Our experience has shown that not more than 20 per cent of applicants will meet the requirements for satisfactory hand-foot coordination for the usual foot-press jobs. The less hand-foot coordination the applicant has, the greater is the accident hazard he presents in the operation of his machine. In a considerable number of cases, we have found operators selected by test who produced at a rate more than 40 per cent greater than the average operator and were, at the same time, accident-free. We suspect, but have not yet sufficient proof, that most of the accidents occurring to operators on hand-foot operated devices

are due to lack of appropriate hand-foot coordination ability, which could easily be ascertained by testing such operators.

The bearing of the foregoing discussion upon the design of machines and fixtures should be apparent. Even the standard left-handed sewing machine might be considerably improved by the addition of certain foot elements to those already employed. Very few power-sewing-machine operators can stop the machine at exactly the point desired and with the needle in the raised or lowered position required for the next adjustment or positioning of the work. Usually the operator must use the right hand to give the handwheel a part turn to put the needle in the proper position. Yet it should be quite possible to effect this mechanically with a suitable pedal for the left foot.

It should also be readily possible to reverse the direction of sewing by such a pedal for the left foot, or by depressing with the right heel the operating pedal now controlled by the right foot, instead of the present awkward control through a lever moved by the right hand at a point near the handwheel. Even the change in position of the present reversing control to a point near the right hand as it follows the work would be a great help in simplifying the operation.

The matter of strength of the operator, particularly on foot presses, should also be taken into account. Some of the older presses require the expenditure of an amount of energy on the part of an average person that makes them unduly fatiguing and that impairs the speed and accuracy of the work. The assignment of heavy-limbed and robust operators to such mechanisms is most desirable. Still more desirable is the solution offered by new power-driven equipment that can be operated more speedily and with a minimum of effort.

On many presses and similar devices, especially on certain boxmaking machinery employing a plunger, the hazard to the operator can be materially lessened by introducing supplementary controls that permit operation only when the hands are clear of the danger area. These can often be introduced without any material change in the cycle of operation and without any sacrifice of production. In fact, their effect in reducing fear of the machine may be beneficial to production in the long run.

Other Implications

An enormous amount of experimentation still remains to be done in this area of hand-foot coordination. Little has been done thus far to measure such coordination as it bears upon the operations of driving an automobile or piloting a plane. Still less has been done to relate it to performance in playing the piano, the pipe organ, or any other instrument employing foot pedals. Yet it is evident from the wide range of this ability that it must have a profound bearing upon success or failure in such performances.

The reason for the lack of experimentation in this area is almost certainly the erroneous belief that the attainment of the necessary level of coordination is almost entirely a matter of practice. Even on the simplest cycles this belief is now shown to be false, whereas on the longer cycles involving complicated dual and bilateral hand elements the range of innate abilities and the attendant handicaps imposed upon some persons by lack of such abilities in suitable degree are shown to be very great.

When a job requires a high level of hand-foot coordination as well as a high level of perceptual ability in controlling such coordination, the number of persons who can turn in a satisfactory performance is indeed small. Lack of musical aptitude in the affective or emotional area may not be the only factor limiting the number of competent musicians; the number may be even more definitely restricted by the scarcity of the coordination ability that is required. This view receives some support from a consideration of the large number of persons who listen to music in proportion to the number who produce it.

The implication for industrial engineering from these considerations is that jobs should not be set up for virtuosos, *i.e.*, for operators who must possess high innate abilities in too many areas of these wide individual differences. High speed in hand operations, plus marked hand-foot coordination ability, plus high speed and accuracy in visual inspection can be found in very few of the applicants who come to the employment office. These facts of individual differences are the best argument for division of labor in terms of the dexterities and perceptions required by the processes.

CHAPTER IV

INSPECTION

Most production employees are required in the course of their work to inspect constantly the product under their control. Most of such inspection is *visual,* but it may be *tactual,* requiring the use of the fingers and hands; *auditory,* involving the sense of hearing; or it may even involve the *olfactory* and *gustatory* senses, relating to smell and taste, respectively. Where inspection ability is called for in the work, an adequate series of aptitude tests for employees must include one or more inspection tests in the sensory areas involved. Where specialized jobs in inspection are set up and employees designated as inspectors devote their full time to the work of inspection, the need for such tests is even more imperative.

The Illusion of Perfect Vision

It was long thought that good eyesight, *i.e.,* high visual acuity, was the best guarantee of satisfactory visual inspection ability and that keen hearing, *i.e.,* superior auditory acuity, was the best guarantee of auditory perceptual ability of a high order. It was pointed out in the Introduction, however, that good eyesight, even superior eyesight, is no guarantee of effectiveness in perception. It may be likewise inferred that high acuity of any sense organ is no assurance that the special type of perception depending on that sensory area will be of a high order of effectiveness. *The physiological and anatomical reasons for this seem to be that the perceptual areas of the brain are not identical with the sensory areas, but lie adjacent to them.* Thus, the visual perceptual areas lie to the right and left of the visual sensory areas in the occipital lobe of the brain. Impairment of these visual perceptual areas is known to interfere with visual perception, whereas the visual sense itself may remain unaffected.

In industrial psychology we are concerned with measures of the visual sense, such as acuity, depth perception, astigmatism, etc., because certain visual sensory defects are directly related to visual perceptual impairment. Even when the measurements show the sense organ to be normal, additional measurements of visual perception are required. One such study[1] indicated that impaired vision materially decreases the speed as well as the accuracy of perception and that, in an experiment involving a quantity of the materials actually inspected by specialized inspectors, both the speed and accuracy of the inspection declined rapidly with increased impairment of vision.

It is exceedingly important that persons on specialized inspection work should have normal vision, or vision corrected to normal. Beyond this, they should have a level of perceptual efficiency established by suitable perceptual tests. A single test of visual perceptual ability may not be adequate for measuring the perception that is sometimes involved in the work.

Kinds of Visual Perceptual Ability

There seem to be at least two kinds of visual perceptual ability. One is the ability to discriminate small differences, such as those involved in the test for inspection of lead pencils for concentricity of leads. The inspection of small mechanical items, such as parts of watches, miniature cameras, etc., involves this type of perception. The other type involves span of perception, *i.e.*, the ability to survey an area, such as that presented by a sheet of tin plate, a sheet of bank-note paper, or a quantity of cloth, to detect small imperfections.

Tests for Discrimination of Small Differences

It is comparatively easy to design a suitable test for the discrimination of small differences. Four such tests of the pencil-and-paper variety are shown in Figs. 3*a*–3*d*, pages 37–40. Each test begins with items showing such marked differences that almost anyone can perceive them. As one works toward the end of the test, checking each imperfect item, he encounters smaller

[1] Holger D. Oleen, Why Pick Inspectors Who Can't Inspect? *Factory Management and Maintenance,* vol. 96, no. 1, pp. 83-85, January, 1938.

and smaller differences, until he reaches an area beyond the limit of his perceptual ability and in which he can only guess. Forms

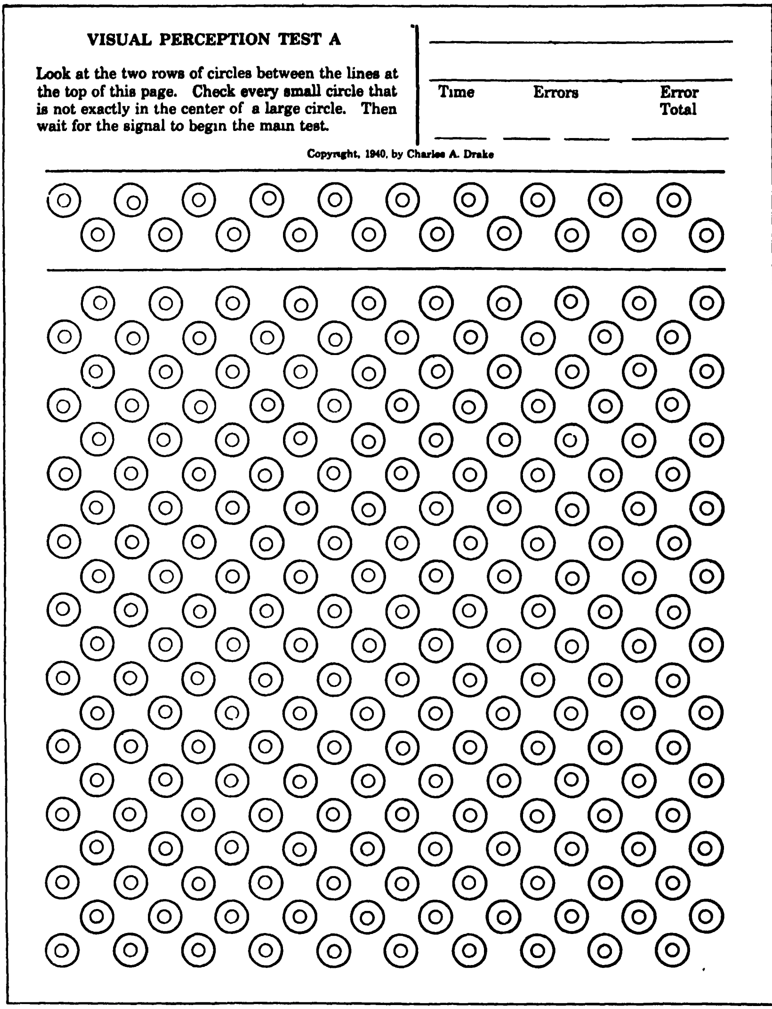

FIG. 3*a*.—Visual Perception Test A.

A and B of this series are applicable to persons in general, while Forms C and D should be given only to individuals who have made exceptionally good scores on Forms A and B. Forms C and D impose an added strain on perceptual ability, because the

separate items are not isolated by a circle and therefore tend to run together.

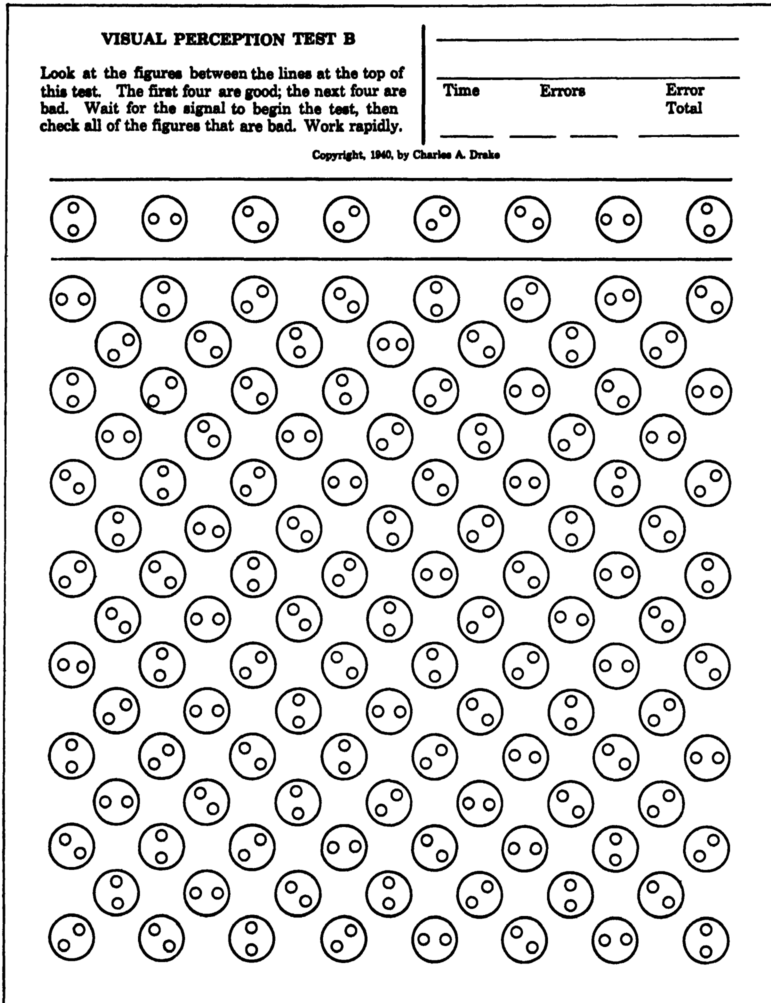

FIG. 3*b*.—Visual Perception Test B.

These tests involve perception in only two dimensions. Where machine parts are to be inspected, the tests must involve objects with three dimensions. Sometimes it is desirable in such a situation to use the actual parts or simulated items that conform

to the perceptual requirements of the job. A few tests of this type have been constructed.

VISUAL PERCEPTION TEST C

Look at the figures between the lines at the top of this test. The first two are good; the next seven are bad. Wait for the signal to begin the test, then check all of the figures that are bad. Work rapidly.

Time	Errors	Error Total

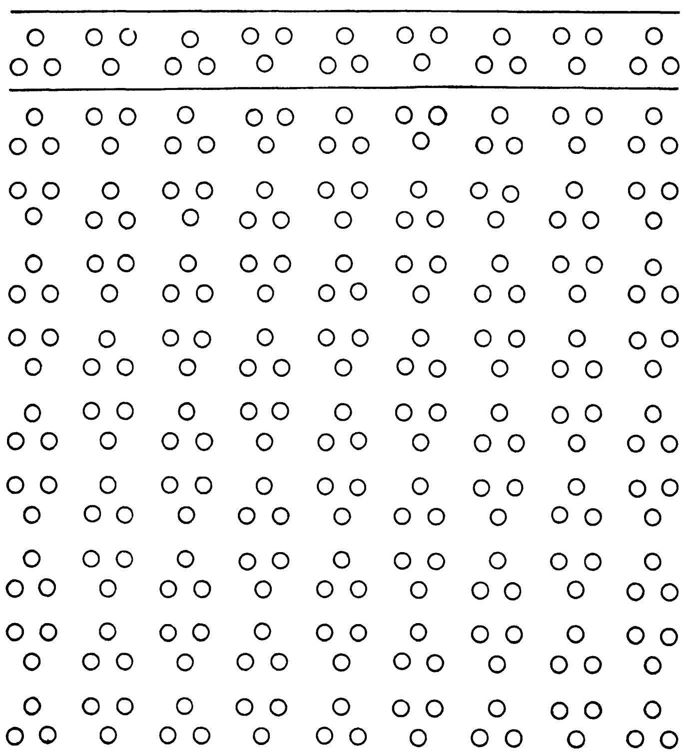

FIG. 3c.—Visual Perception Test C.

In the work for the Eagle Pencil Company, one inspection test of this sort made use of the small aluminum spirals used in mechanical pencils. This is shown in Fig. 4, page 41. Fifty standard spirals bearing a punch mark 2½ turns from one end

were mixed with 50 off-standard spirals punched at distances other than 2½ turns from the end. The testing procedure in-

VISUAL PERCEPTION TEST D

Look at the figures between the lines at the top of this test. The first two are good; the next five are bad. Wait for the signal to begin the test, then check all of the figures that are bad. Work rapidly.

Time	Errors	Error Total

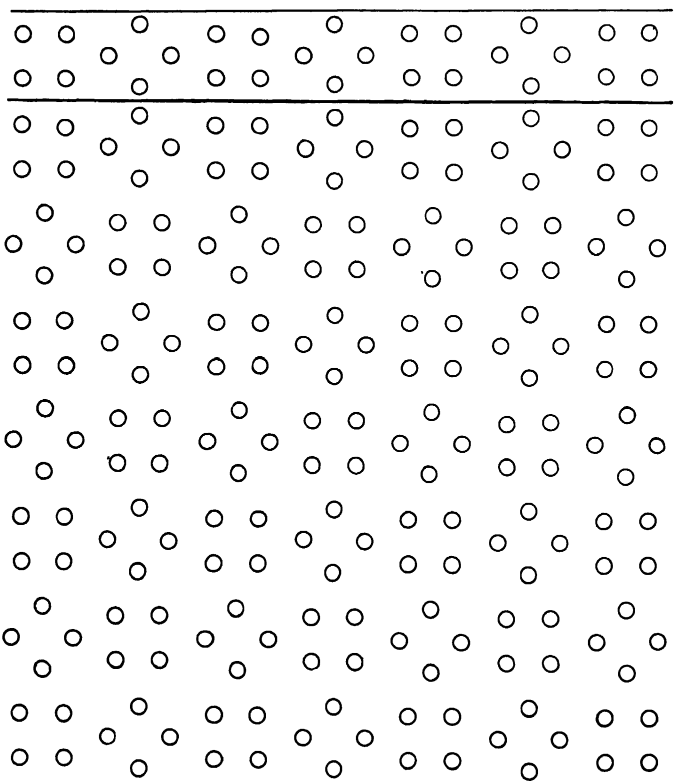

FIG. 3*d*.—Visual Perception Test D.

volved the separation of the standard spirals from the off-standard as quickly as possible. The test was scored for both speed and accuracy.

The second test of this sort constructed for the same company

made use of 120 metal pencil cases in six different colors. Thirty of these cases were rendered defective by punch marks. The

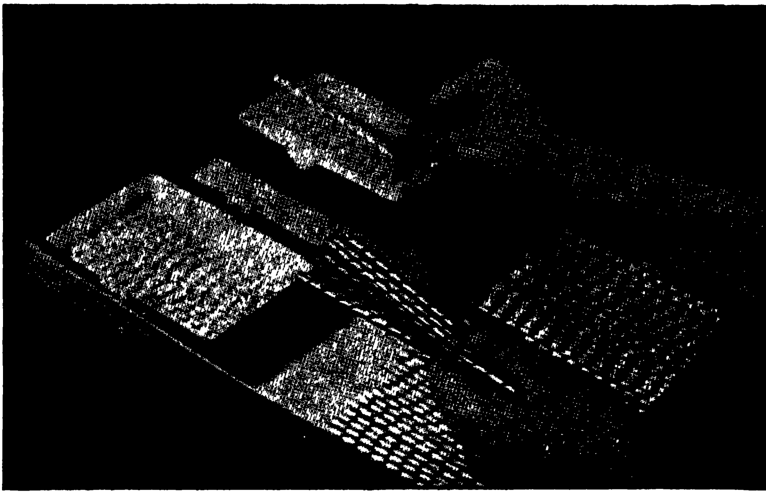

Fig. 4.—Spiral Inspection Test.

task was to sort the good cases into six compartments by color and the defective cases into a seventh compartment, as shown

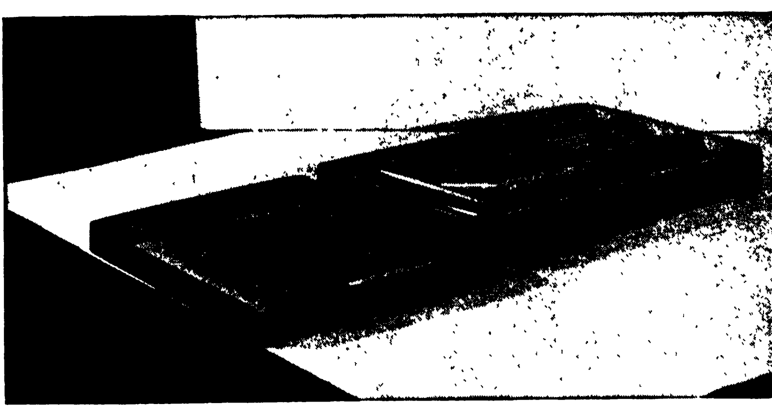

Fig. 5.—Case Inspection Test.

in Fig. 5, above. This latter test involved a factor not present in the Spiral Inspection Test, namely, that of sorting the standard pencil cases.

Neither the time scores nor the error scores on these tests are significantly related to the dexterity tests, as shown in Table II, page 27. All the error scores on both tests are so near zero

TABLE IVa.—SPIRAL INSPECTION TEST—TIME SCORES

Time, Minutes	Frequency
2.50– 3.49	4
3.50– 4.49	21
4.50– 5.49	48
5.50– 6.49	83
6.50– 7.49	63
7.50– 8.49	28
8.50– 9.49	17
9.50–10.49	10
10.50–11.49	5
11.50–12.49	1
12.50–13.49	2
13.50–14.49	1
14.50–15.49	1
	284

TABLE IVb.—SPIRAL INSPECTION TEST—ERROR SCORES

Errors	Frequencies		
	Applicants	Operators	Retest—both
0	17	2	8
1	29	8	10
2	19	3	5
3	17	4	5
4	14	3	1
5	4	2	2
6	7	1	3
7	4	2	2
8	5	0	3
9	7	1	1
10 and up	12	14	6
Totals...	135	40	46

that they are clearly not related to the manual skills measured by the other tests. The low relationship between the time scores on these tests and the scores on the manual tests is, of course, due to the manual manipulation called for in the sorting of the

test items. Both of these tests showed wide ranges of scores in both time and errors, as shown in Tables IV*a* and IV*b* and V*a*

TABLE V*a*.—CASE INSPECTION TEST—TIME SCORES

Time, Minutes	Frequency
3.00– 3.49	4
3.50– 3.99	5
4.00– 4.49	13
4.50– 4.99	38
5.00– 5.49	52
5.50– 5.99	67
6.00– 6.49	49
6.50– 6.99	46
7.00– 7.49	25
7.50– 7.99	15
8.00– 8.49	13
8.50– 8.99	5
9.00– 9.49	2
9.50– 9.99	2
10.00–10.49	1
10.50–10.99	2
	339

TABLE V*b*.—CASE INSPECTION TEST—ERROR SCORES

Errors	Frequencies		
	Applicants	Operators	Retest—both
0	138	12	15
1	108	9	14
2	27	9	11
3	34	5	5
4	21	2	1
5	6	1	
6	2		
7	4	2	
8	3		
9	2		
10			
11	1		
Totals..	346	40	46

and V*b*, pages 42 and 43. Applicants and experienced inspectors both showed about the same distribution of these scores. We

should not be surprised at the poor showing of the experienced operators, since they were selected by interview and not tested for the ability called for in inspection. A retest of 46 of the operators and employed applicants on the Case Inspection Test after three months, indicated in the last column of Table V*b*, shows little improvement in the error scores. An examination of the improvement in the time scores of these 46, not tabulated here, shows an average gain of less than one minute, owing chiefly to the reduction among those whose first time was excessively long. These results indicate that the test is measuring something innate—something not materially affected by training and experience. Tests of this sort generally are not so reliable as tests of motor ability. Reliability is usually sacrificed in the interest of keeping the test within reasonable time limits. It is a well-established fact that reliability can be increased by lengthening the test, and it is possible to predict by a suitable statistical formula the amount of increase necessary to attain any desired reliability.

The Critical Score

In the industrial application of inspection tests, we are primarily concerned with the identification of the persons who lack the abilities called for on the job. If the persons who lack these abilities can be eliminated by suitable tests, our training procedure and performance on the job will have a high probability of success. We are concerned, therefore, with the establishment of a *critical score* for speed as well as for accuracy. Applicants scoring beyond the points designated will not be chosen for the type of work involved. (See also Chap. XIII, page 126.)

Test for Span of Perception

The design of a suitable test for span of perception is difficult but not impossible. It is certain to be expensive, and will consequently not be done except on a consulting basis for an industry sufficiently large to justify the expenditure. Probably the most urgent need for such a test is that of the companies producing tin plate, each sheet of which must be individually inspected before packaging. Because of the size of such sheets, and consequently the necessity for perception over an area much larger

than that of the usual printed test, we may infer that a pencil-and-paper test will not be suitable for this purpose.

One should not expect to find a test appropriate for measuring the ability to detect small differences adequate for the selection of personnel who are called upon to inspect on jobs involving span of perception. The writer has specifically warned against the use of Visual Perception Tests A to D in such situations.

Intelligence and Inspection Ability

In several instances known to the writer, companies have endeavored to select inspectors by the use of intelligence tests, clerical aptitude tests, and other similar instruments. These efforts have failed, somewhat to the surprise of the experimenters. Two investigations by the writer, using Visual Perception Tests A and B, have shown that both the time scores and the error scores on these tests are not significantly related to scores on the Henmon-Nelson Test of Mental Ability or the Otis Self-administering Test of Mental Ability. It may be inferred, therefore, that intelligence as measured by intelligence tests, academic scholarship, high-school graduation, and the like would not be helpful in the selection of inspectors for industrial work.

Illumination and Background

The matter of illumination of the items to be inspected and of the background against which they are inspected is of extreme importance. It must be borne in mind that there is a determinable optimum illumination for each inspector on each job. Most inspection jobs are poorly set up with respect to both source and amount of light, and with respect to the color and other qualities of the background. Some preliminary experimental work has shown that in all surface inspection, as well as in the inspection of transparent and translucent objects, polarized light makes for better perception of small differences.

Periodic Checkups

Deterioration of both vision and visual perceptual ability is known to come at varying speeds with the advance of age. Farsightedness seems to interfere with inspection ability far more than nearsightedness on most jobs. It is important, therefore, to make repeated checks on both vision and visual perceptual ability of the personnel on specialized jobs. This is only slightly less important with operators who must inspect while they produce or assemble the product. In either situation, such a check should be made at least once, and preferably twice, each year. Where inspection is found to be unsatisfactory, such a check should be made immediately as the first step toward correction of the situation.

Testing of Other Perceptual Areas

What has been said of visual inspection is equally true of inspection in other sense areas. Inspectors of glass bottles and other products often "clink" the units together, listening for the sound in the detection of imperfections. In high-grade products, such as crystal glassware, such a test is of the utmost significance. The writer knows of one instance in which psychologists applied a battery of a dozen or more tests for the selection of such inspectors, but included in the battery not a single test of auditory perception or even of auditory acuity!

It is entirely possible to design a test of auditory perception, although the design is somewhat complicated by the fact that there are at least three dimensions of auditory perception, each of which may need to be tested separately. *Pitch* of sound constitutes one dimension of measurement, *loudness* another, and the third, and probably most significant for industrial uses, is *consonance*, or purity of tone. The recognition that a piece of glassware is cracked seems to depend primarily upon perceptual ability of the last-named sort.

Certain peculiarities of hearing, such as tone deafness, tonal islands, and others, may incapacitate a person for work involving auditory perception. Defects of the sense organ itself have,

in all probability, effects on auditory perception analogous to those found in visual perception.

Industrial applications of measures of perception in taste, smell, and touch are merely less common but not less important than those of sight and hearing. It is quite possible to design suitable tests of perception in these areas, following the principles that apply to the design of tests in the visual area.

Inspection and Quality Control

There is no situation in industry where a suitable test will show more striking results in selection than in the measurement of perception. Both speed and accuracy of inspection can be enormously increased in the average situation by the selection of competent personnel through the use of suitable tests. Quality control in industry is absolutely dependent, in the final analysis, upon the personnel making the last check on the product.

It is not uncommon to find an unnecessarily high number of imperfect items produced where the operators know that later inspection is inadequate or incompetent. Rigid inspection of the product thus has a redundant effect on the production process and tends to maintain a high level of quality. This can be guaranteed only by the positive assurance that the final inspectors are competent, as a result of selection by appropriate tests.

CHAPTER V

PERCEPTION

It is a rather curious fact that experimental psychology made a quick jump in its interests and efforts from sensory testing to the measurement of intelligence, with scant attention to perception. Logically the measurements in the several sensory areas should have been followed by measurements in the perceptual areas associated with, and dependent upon, those sensory areas. It is difficult to explain this oversight, but it may be suggested that experimentation was diverted to illusions and the other more spectacular manifestations of perception, whereas the lines of investigation that might have made direct contributions to the selection of industrial personnel were largely ignored.

Our attention was attracted to the field many years ago by empirical observation of the fact that many students in laboratory courses in zoology and botany could not see what they were supposed to see under their microscopes. The obvious inference from the condition noted was that these students were suffering from eye defects and needed glasses. Yet careful examinations of the eyes of these handicapped students, made for the purpose of fitting suitable lenses, disclosed no defects in the eyes themselves and even, in some instances, indicated quite superior visual acuity.

One may argue that, since what is perceived is the result of a learning process in the course of growth from infancy to maturity, these students had not learned to see what they were supposed to see. This argument has little weight when it is found that despite good teaching, and intensive effort on the part of the student, the function still cannot be performed. We are forced to the conclusion that defect in some structure beyond the eye itself is responsible.

On the other hand, certain defects in the structure of the visual organ may enhance perceptual ability. Thus it is ob-

served that some color-weak and color-blind individuals can perceive items that are not perceived by other persons of normal vision. It is interesting to note in this connection the recent announcement by the press that certain color-blind individuals in an air force are able to perceive camouflaged constructions that are imperceptible to others with normal eyesight.

Early Experimental Work

When we found, in the course of industrial experimentation, that persons of normal vision often could not perform satisfactory inspection operations, we undertook the development of tests for this function. The first results gave such wide ranges of scores in both time and accuracy for persons of normal and superior vision, and, generally, poorer scores for those of impaired vision, that we made this a major field of investigation.

An error in our early work was the tacit assumption, following textbook treatments, that industrial inspection involved reasoning processes primarily, with perception as an intermediate process between sensation and reasoning. This view we no longer hold. Our present view is that acceptance or rejection of an item being inspected is based upon a perceptual process alone and that reason is not involved unless the inspector is called upon to explain *why* he reacted in a given way.

It is considered an unfortunate handicap in the attempts to measure special areas of perception that the experimenters expect their measures to be significantly related to intelligence, *i.e.*, to intelligence as measured by the usual intelligence tests. The usual high intercorrelations of tests of "social intelligence," "sales aptitude," "legal aptitude," and others of the kind with tests of general intelligence should be a warning. Any considerable overlapping of this kind merely indicates that the special test is measuring to a substantial degree the same thing that is measured by the general test.

Applications of Perceptual Facts

There are obviously many important industrial and other applications of the known facts of perception. It is well understood that certain persons have superior ability to see forms at

night or under other conditions of reduced illumination. This ability to react to reduced stimuli, or "minimal cues," is a great asset in driving motorized equipment at night, in fog, and under similar adverse conditions. It is also known that a diet deficient in certain vitamins may result in impaired vision and visual perception—impairment that may be corrected by suitable dietary changes. Certain colors can be perceived at greater distances than other colors of the same intensity, and some colors have such meanings as "danger," "caution," etc., associated with them through previous training and experience.

A practical "intuitive psychologist" of the writer's acquaintance had experienced several collisions in which other cars had struck his own from the side. He then painted the wheels of his car a brilliant vermillion and thereafter claimed exceptional freedom from collisions and "close shaves." He argued that a bright color, especially red, is perceived more quickly and meaningfully than black or gray.

Social Intelligence—An Aspect of Perception

We suspect, from our results, that much that passes for intuition and "hunches" should properly be called perception. Thus we have observed persons possessed of a high order of "social intelligence" who were quite at a loss to explain convincingly the reasons for their right conduct. The answer probably lies in their superior social perception, their ability to react to a complex of social cues imperceptible to others less well endowed.

Such individuals have enjoyed great advantages in our society in certain lines of effort. We have placed premiums upon their ability to control the behavior of others—control based upon their superior insight or perception of human interrelationships. Often this ability is possessed in a high degree by individuals whose general intelligence, as measured by the usual tests, is only average, or even below average.

This so-called "social intelligence" certainly involves the ability to receive and interpret a variety of small cues in the visual, auditory, kinesthetic, and perhaps other sensory areas. It also must involve some specialized ability in the perception of relationships among the factors involved in a complex social situation. This latter factor would seem similar in kind to, but not

identical with, certain perceptual activities in inspection, since responses are direct, immediate, and devoid of other characteristics that indicate reflection or reasoning on the part of the person making them.

We have been further interested to observe individuals of high general intelligence who could not, even under expert instruction, learn to play a game such as bridge with real skill. They seemed to lack a "card sense"—actually what is probably a special form of card perception—whereas others less intelligent quickly developed a high ability, often through practice and experience without instruction.

This last inference agrees with what is often observed in industrial inspection. New inspectors, selected by test, soon become as speedy and as accurate as those of long experience, even without specific training. Their measured perceptual ability is quickly realized in production on the job. We have failed in our attempts to make even mediocre inspectors of persons who made poor scores on the tests.

Salesmanship and Perception

The field of salesmanship is one in which specialized perceptions must be of great importance. Many sales managers maintain that good salesmen are born, not made. Certainly the standard techniques of selling, like the rules of bridge, can be learned by almost anyone. But the perception of the situation that determines the applicability of the techniques depends upon some innate ability. Such ability is probably, like other perceptual abilities, little subject to change by training.

The problem of selecting salesmen, therefore, becomes one of selecting persons of high measured sales perception. This would seem to be simple. There are several tests of "salesmanship" available. Why is the success of these tests so moderate in actually selecting good salesmen? It is our view that these tests are not measuring sales perception and therefore are not giving a measure of the function upon which success depends.

The trouble with such tests, and with certain others—such as the writer's test of bridge ability, which he finally abandoned-is that they are verbal. That is, they are expressed verbally—in words—and the responses are verbal. The sales situations

are presented through the medium of words and sentences. Perception of words and verbalized ideas is not the equivalent of sales perception but only the equivalent of itself—perception of words and sentences.

It will not be simple to devise adequate tests of sales perception. Tests with written questions and answers will not serve. Nonverbal tests must be developed. This process will be both long and costly. We believe the clue lies in our own nonverbal tests of visual perception, but no private individual will undertake the further expensive research required. If a highly successful test were devised, it would be quickly duplicated, paralleled, and appropriated by those persons in need of such a measure, and the original investigator would fail to recover his costs. Such development will not be undertaken, therefore, except by a foundation or by a trade association.

The obstacle to successful development of the required instruments is, in our present view, the lack of techniques in the measurement of nonverbal perceptions. While it would be scientifically presumptuous to attempt to lay down definite principles to guide such a project, the preliminary lines of experimentation are fairly clear. Certainly the trait to be measured is more easily and directly accessible than accident-proneness, already revealed by a simple technique (see Chap. XII).

The economic and social consequences of the successful development of a sales aptitude test can scarcely be overestimated. Even moderate success, from a statistical point of view, would greatly reduce the high costs of salesman training and turnover as well as the emotional stress and strain in the personnel, particularly in those who would otherwise fail.

Engineering Aptitude and Perception

In the field of engineering aptitude, the experimenters have devoted their efforts very largely to perceptual measurements. Such measurements have to do with relationships, particularly with structural relationships in three dimensions of space. It is our view that such efforts will meet with only limited success as long as they are confined to static or geometrical relationships. The requirement in this field seems to be for dynamic

tests, or tests of constantly changing relationships, to which the calculus rather than geometry is more appropriately applied.

Individual Differences in Perception

The observation that leads to the foregoing inference is that, as an individual is forced to increase the speed of his perceptual reactions in inspection, the accuracy of such reactions decreases. A similar phenomenon is observed when an individual is required to increase the speed of motor reactions in a long or complex cycle that does not involve much perception. In the latter situation errors and mistakes increase rapidly with the increase in speed, finally resulting only in frantic random reactions. In both areas there are wide individual differences that imply similar differences in individual effectiveness.

There are still other individual differences in perceptual abilities that must be taken into account. We have observed in the extensive data collected on Forms A and B of the Visual Perception Tests a pronounced tendency to bimodality in the distributions. That is, there are two points of greatest concentration of scores. This phenomenon is most pronounced and persistent in the time scores, but it is also observed in less degree in the error scores. The only plausible hypothesis we can offer for this is that there are two types of reaction in this area of perception.

Introspective evidence suggests that some persons make a "total" or unitary reaction, while others make a series of analytic reactions to each test item. The former would react perceptually to each item as symmetrical or nonsymmetrical; the latter would examine each of the several relations of one circle to the others—such as vertical distances and horizontal distances—and would, presumably, give a slower reaction. The foregoing hypothesis seems worthy of further investigation as to its practical diagnostic significance in choosing inspectors. Apparently the unitary type reaction is quicker, and there is reason for suspecting that it is more accurate.

Other Factors Influencing Perception

The effects of fatigue, alcohol, drugs, and emotional states upon perceptual speed and accuracy constitute an area of investigation that should be further explored. We have observed wide individual differences in the breaking down of perceptual efficiency under the influence of fatigue and alcohol. There is reason to suspect that accident-proneness is increased when the level of perceptual efficiency under such influences breaks down more rapidly than the level of motor reaction. But in some persons the motor reactions deteriorate more rapidly than the perceptual reactions, thus leading to the inference that such persons do not become accident-prone under these influences either so rapidly or so completely as the others.

Perception and Quality Control

The intimate bearing of the individual's perceptual reactions upon quality control in industrial production must be quite apparent. What holds for the visual area must be equally true of other areas of perception. The analogous deterioration of kinesthetic and auditory perceptions should have similar results in poorer work and increased hazard to the workers. Consequently, both the innate perceptual characteristics of employees, particularly of inspectors, and their sets of personal habits should be matters of paramount importance in selection and placement.

Two Perception Tests

A test designed to measure tactual, or touch, perception is shown in Figs. 6*a* and 6*b*, page 55. This test consists of seven cylinders, each covered with a different grade of sandpaper. The cylinders are arranged in a random order before testing. The testee is then required to rearrange the cylinders in order from coarsest to finest by touch alone, the cylinders being obscured by a curtain. Scores are in terms of total time required and the number of errors in the arrangement.

The results from applications of this test indicated that there were too few items and that the differences in degrees of fine-

ness between items were too large and therefore too easily perceived. Such a test should probably consist of at least 20 items.

Figures 7*a* and 7*b*, page 56. represent a test designed to

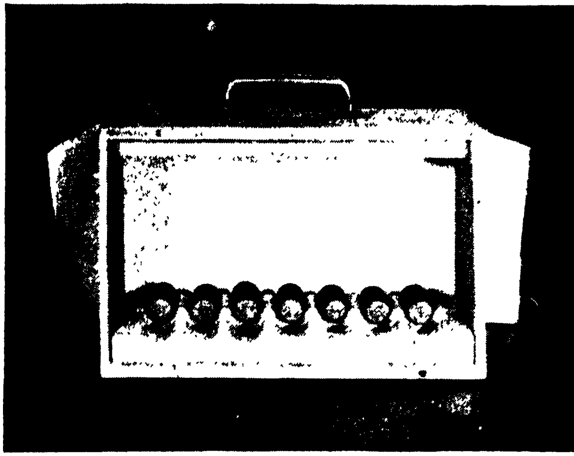

FIG. 6*a*.—Tactual Perception Test.

measure a certain aspect of kinesthetic, or muscular, perception. The testee is required to place a small pin through the sloping eye of a larger pin without the aid of vision. The testee manipu-

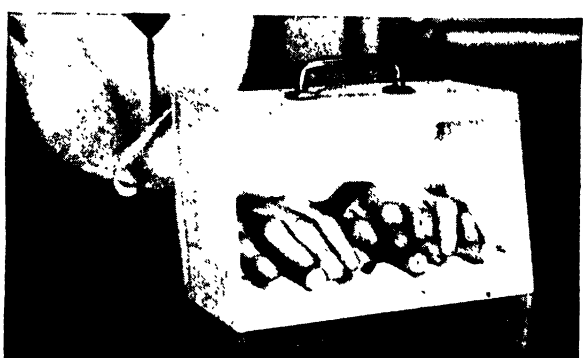

FIG. 6*b*.—Tactual Perception Test.

lates the large pin with the left hand, while he "feels" for the hole by attempting to insert the small pin with his right hand. This preliminary exploration is performed with the small pin

in the first hole, the one nearest the testee. Having determined the angle of the eye in the large pin, he then removes the small pin and places it in the one of the three holes further to the

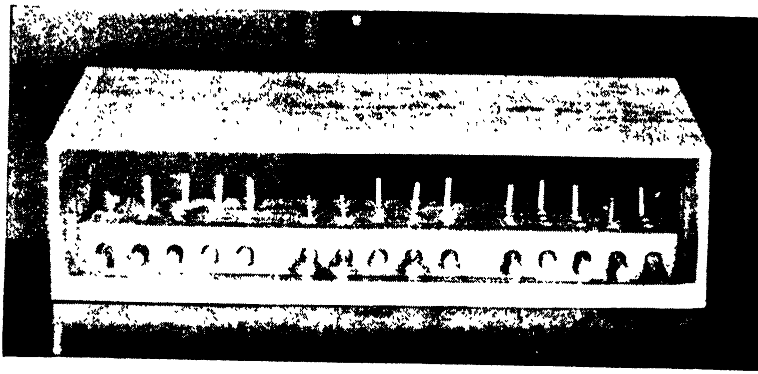

FIG. 7a.—Kinesthetic Perception Test.

rear that slopes at the angle corresponding to the angle of the eye. He slides the large pin into the block until the eye is in line with the small pin, and then pins it fast by pushing the

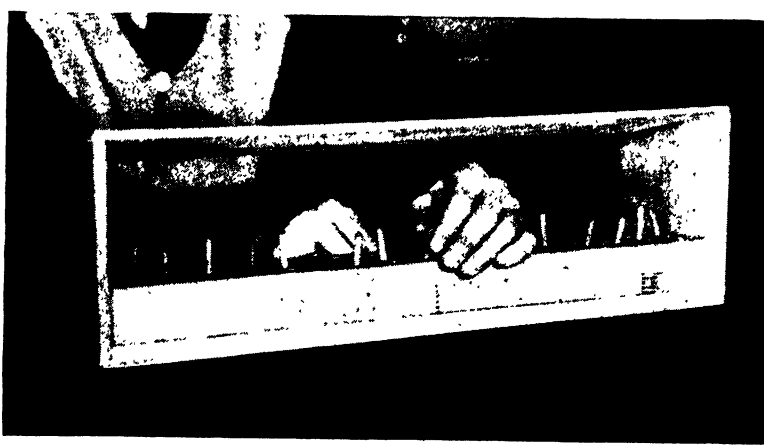

FIG. 7b.—Kinesthetic Perception Test.

small pin downward through the eye. He then goes on to the next item to the right of the first. Scores are in terms of time required only, since the testee is obliged to complete each item without error before going on to the next item.

The reliability of this test was impaired by the fact that its design violated one of the principles of perception testing. All items were of equal difficulty, whereas it should have presented the testee with items of increasing difficulty in performance. It should be comparatively easy to design a test of kinesthetic perception which conforms to this principle.

Either of the foregoing tests may be motor-driven to afford definite time intervals within which perceptual adjustments must be made. The desirability of designing such tests in motor-driven form would, however, have to be determined in relation to the particular problem being studied.

Where the individual is himself in motion through space, as in driving an automobile or piloting a plane, the time factor must be added to the perceptual testing procedure. The economic and safety reasons for this are evident. If perceptual reactions tend to break down as the rapidity of the stimulus increases, the point at which this adds materially to the accident hazard must be determined. In other words, if the individual can perceive and react adequately only up to a speed of 100 or 200 miles per hour, he should not be required to attempt adjustment and control at higher speeds. The present evidence is to the effect that he cannot, by practice and instruction, materially raise his upper limit of this ability.

Apparently there are at least two types of individual differences involved in such situations. One is the speed of perception itself, with its accompanying motor reaction or inhibition of reaction (such as to turn the steering wheel or to hold it as it is); the other is the individual's ability to react to reduced cues or slight stimuli (as in fog, rain, diminished illumination, and noise). Both of these should be measured, and they must be measured in any complete inventory of aptitudes and abilities.

A Complex Perceptual Area

The extreme complexity of some perceptual reactions is indicated in the following analysis of "Hand," as the term is used in the textile industry. The need for clear definition of the terminology is indicated in the explanatory paragraph and table reproduced below.[1] From the standpoint of test construction,

[1] *Textile World,* vol. 91, no. 1, p. 81, January, 1941.

it is clear that "Hand" involves a combination of tactual, kinesthetic, and visual perceptions, and that each of these perceptual areas will require a separate test. The minimum number of tests will, therefore, be three, but the characteristics of the kinesthetic perceptions may require the design of several quite different kinesthetic tests.

For consistency and better understanding among those dealing with the hand of fabrics, representatives of the American Association of Textile Chemists & Colorists and of Committee D13 at a joint meeting proposed that the terms in the accompanying table be used regularly in naming the physical properties of fabrics related to hand and for describing the corresponding components of hand. Simple words understandable to the layman, as well as to the scientist, were chosen. As research progresses on the evaluation of hand and as standard tests for each property are recommended, definitions of the properties measured by the tests will be added to the table.

Physical property	Explanatory phrase	Terms to be used in describing the range of the corresponding component of hand
Flexibility.........	Ease of bending	Pliable (high) to stiff (low)
Compressibility....	Ease of squeezing	Soft (high) to hard (low)
Extensibility.......	Ease of stretching	Stretchy (high) to non-stretchy (low)
Resilience.........	Ease of recovery from deformation in which rate of recovery is without limits and thus includes elasticity (instant recovery)	Springy (high) to limp (low). Resilience may be flexural, compressional, extensional, or torsional
Density...........	Weight per unit volume (based upon A.S.T.M. standard measurement of thickness and fabric weight)	Compact (high) to open (low)
Surface contour....	Divergence of the surface from planeness	Rough (high) to smooth (low)
Surface friction.....	Resistance to slipping offered by the surface	Harsh (high) to slippery (low)
Thermal character..	Apparent difference in temperature of the fabric and the skin of the observer touching it	Cool (high) to warm (low)

Implications of Perception

It must now be apparent to any thoughtful person that the field of perception is one that deserves immediate and thorough exploration. So much of social and economic well-being is tied up with its peculiarities that these relationships should be measured and understood. This is not an area inaccessible to measurement. It is, on the contrary, easily accessible to the experimenter who will abandon some of his erroneous preconceived notions and start afresh with the findings of current experimentation.

CHAPTER VI

DUAL COORDINATION

If an assembly job can be done with one hand, using the other hand or some mechanical aid for holding the work, it is generally believed that greater economy will result from the simultaneous assembly of two units of the product, using both hands. Whereas this belief is generally justified, it must be accepted with caution in particular instances. Both the nature and extent of the assembly motions and the fact of wide individual differences in the operators are variable factors which influence the results.

Time- and motion-study specialists have long contended that dual operation on assembly should effect, roughly, a 30 per cent saving in time per unit of product. They also concede that such savings may be somewhat less where the operation cycle is short. Observations both of jobs and experimental tests support their views when the persons under observation or test are "average" operators. But it is possible to utilize the fact of individual differences to select, by suitable tests, certain operators who can achieve a saving as high as 55 per cent.

Dual Operation Ability—A Specialized Function

In the process of making psychological job analyses as a basis for test specifications and designs for tests, it was observed that numerous dual operations were dual in name only. The operators were in fact alternating the hand motions that were supposed to be done simultaneously. Later, when these operators were tested on the dual tests, it was found that many of them were inherently incapable of achieving dual coordination. Even more striking was the discovery that, after intensive training under close supervision in a training department for two weeks, they were still incapable of attaining a performance that warranted the time and effort spent in training.

It is now fairly clear that dual operation ability is as specialized and as rare as inspection ability. Test results indicate that not more than 15 satisfactory dual operators are found among 100 applicants for employment, and about the same number among 100 qualify for training for inspection. If dual coordination and inspection ability are called for in the same operator, the combination will be found in about only 7 to 10 of the 100 applicants.

Many operators of high potentiality for dual operation are to be found among old employees who have been doing nondual work. There is no evidence that the trait is modified by nonuse, since the experienced operators show the same distribution of test scores as the new applicants. Also, a good dual performer on one test tends to be, but is not necessarily, a good performer on other dual tests.

Tests of Dual Coordination

The tests used in the measurement of dual ability were Pin Board E, Fig. 8, below, the Controlled Turning Test, Fig. 9,

Fig. 8.—Pin Board E.

page 62, and the Left-right Turning Test, Fig. 10, page 63. In each case the percentage saving was computed by dividing the time taken to do the test with the leading hand—usually the right—into the difference between that time and the time taken to do the test by dual operation. Thus, if the time taken to do the test with the right hand alone was 100 seconds and

the time taken by dual operation was 60 seconds, the savings would be $\frac{100 - 60}{100}$, or 40 per cent.

The three tests used differed from each other in several im-

FIG. 9.—Controlled Turning Test.

portant particulars. The Pin Board involved the placing of 32 pins in each board, the cycle of operation on each pin consisting of six elements. The Controlled Turning Test involved the insertion of 20 slotted keys through holes in a vertical steel plate. It had a seven-element cycle. The Left-right Turning Test re-

quired that long screws be turned, left with the left hand, right with the right hand, through a vertical plate. Its cycle consisted of three elements. In addition to the foregoing differences, it is worth noting that the Pin Board required the manipulation of objects $\frac{1}{8}$ inch in diameter, involving finger dexterity and whole-arm motion. The Controlled Turning Test used ob-

FIG. 10.—Left-right Turning Test.

jects $\frac{1}{2}$ inch in diameter, involving much finger and wrist motion, with two positionings for each unit. The Left-right Turning Test could be performed with either finger or wrist motion predominating, the person tested choosing his own style, while the size of the object, $\frac{1}{4}$ inch in diameter, was relatively unimportant.

The results in terms of savings by dual operation on the three tests are shown in Table VI, page 64. Comparison of the last three columns shows that savings on the Pin Board and the Left-right Turning Test were approximately equal and averaged around 31 per cent, thus supporting the contentions of the time- and motion-study specialists. On the Controlled Turning Test,

however, the average savings are about 11 per cent, with many cases showing negligible savings or actual losses. This result with the latter test is attributed to the two difficult positionings of the keys and the need for greater precision of movement in all the elements of the cycle, as well as the longer cycle itself.

TABLE VI.—SAVINGS BY DUAL OPERATION

Savings, per cent	Frequencies for each test		
	Pin Board	Controlled Turning	Left-right Turning
51 to 55			2
46 to 50	1		9
41 to 45	18		16
36 to 40	49		16
31 to 35	72	2	19
26 to 30	72	10	14
21 to 25	38	12	15
16 to 20	17	15	7
11 to 15	5	22	9
6 to 10	1	25	8
1 to 5	1	13	
− 4 to 0	1	13	2
− 9 to −5		4	
−14 to −10		2	1
−19 to −15			1
−24 to −20		2	1
Totals........	275	120	120

The reliabilities of these three tests are all over .90. That is, successive applications of the tests to the same persons yield almost the same time scores, so that the results may be accepted with confidence in their stability. If the results of the tests can be related to performance on jobs requiring dual operation, their use is justified in preference to the old method of trial and error in selecting dual operators.

The best data for comparative purposes came from the results of the daily records of performance in the training department. Table VII, page 65, shows the comparison in terms of per cent of

"normal" time allowed for three of the dual operations on small-parts assembly.

TABLE VII

Class of operators	Average time at start, per cent of standard	Average time at end (two weeks), per cent of standard	Average gain, per cent of standard
By those who passed tests..	80	97	17
By those who failed tests..	76	91	15

This merely represents a comparison of those who "passed" the tests by scoring in the upper half of the distribution of scores with those who "failed" the tests by scoring in the lower half. When the comparison is made between those recommended for dual operation and those rejected, the results are more striking, as shown in Table VIII, below.

TABLE VIII

Class of operators	Average time at start, per cent of standard	Average time at end (two weeks), per cent of standard	Average gain, per cent of standard
By those recommended ..	103	119	16
By those not recommended.	79	93	14

Whereas the difference in average gains between the two groups is not marked, the comparison between the points at which they begin and finish is startling. The recommended group starts its training above normal, and at a level that the nonrecommended group cannot attain after two weeks of intensive training! Clearly, these results warrant careful selection of dual operators with the aid of suitable tests.

A comparison of the nature of the work on the several dual jobs with the results of the test indicates that the greatest handicap to increased savings through dual operation comes from difficulties in positioning the parts manipulated. When this positioning is difficult, the dual operation usually breaks

down into two separate operations in which attention goes first to the one and then to the other of these operations. When this occurs there is no gain over a one-hand operation in this element of the cycle, since one hand is at rest, or at least ineffective, while the other hand is working.

Mechanical Aids in Dual Operation

Experience points to the fact that certain mechanical aids may contribute largely to the efficiency of dual operation. Better arrangement of parts will cut down the distance of transport by the hands and may facilitate the grasp of parts. This is especially true where the parts can be presented in hoppers or otherwise arranged all in the same relative position, thus requiring a minimum of attention to the element of grasp or pickup. Flared or funnel guides may frequently be used to aid positioning, so that delicate and exact finger adjustments are not required and strain on both attention and eyesight may be minimized.

Savings by Dual Operation

If the dual job is properly set up, and operators are selected by suitably designed tests, it should be relatively easy to raise the average savings by dual operation from the 30 per cent estimated by time- and motion-study specialists to 40 per cent, or, in the case of better-than-average operators, to an even higher figure. Certainly there exists among the applicants for factory jobs a sufficient number of highly gifted dual operators to fill all the jobs that can usually be set up for dual operation.

Random selection, without the use of suitable tests, will lead to the attempt to get dual savings from operators not equipped by nature to effect such savings. Then the method will be blamed, or the mechanical fixtures held faulty, whereas the real difficulty resides in the fact of an unlucky guess about the operators' innate ability. It has long been recognized that people differ in height, weight, color, intelligence, vision, etc. Now it is demonstrated, by the demands of industrial adjustment, that they differ likewise in dual coordination, in hand-foot coordination, and in the ability to combine dual operation with foot coordination.

CHAPTER VII

A MACHINE-OPERATOR TEST

This test, shown in Fig. 11, below, was designed and first applied during the summer of 1940. It represents one of the first attempts to construct a motor-driven test for a machine-

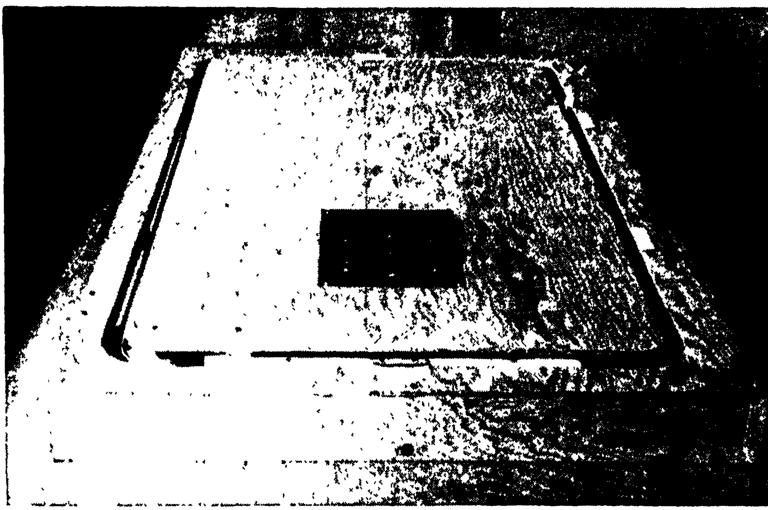

FIG. 11.—Machine-operator Test.

operator job. It has admirably served the purpose for which it was intended and has thereby indicated a line of test development of great significance for the future.

The machines served by these operators transformed materials fed from six different rolls at six different points in the machine into the product, which was packed from one end of the machine by specialized inspector-packers. The machine-operators were required to replace the rolls of material as they became exhausted, anticipating such action by observation of the diminishing rolls. They were also required to make minor

adjustments to the machine, to reduce occasional jams, and to inspect the product from time to time for quality and conformity to standards.

Observation of the job and discussion with the operators and packers led to the conclusion that the essential requirement for successful performance was to keep the machine in production. Some machines in the hands of poor operators suffered frequent stoppages due to failure to replace a roll of material in time, inattention to developing jams, and other causes that were adequately handled by other operators.

No record had been kept of stoppages on each of the dozen machines, of the lengths of such stoppages, or of the reasons therefor. It was consequently not possible to design a test directly from the preventive or prophylactic point of view, as would have been possible if such records had been available. Even if the records had been kept, there would have been no real assurance that the figures were accurate, owing to personal factors such as changes in the records (to "cover up" a poor operator) or incompetent analyses of the causes of stoppages. It was assumed to be better technique, under the circumstances, to make a direct analysis of the job requirements and to design a test to measure the abilities that seemed to be required.

Operation of the Test

The test as designed required observation of six points, paralleling the job, with a set of changing relationships among these six points over a fixed period of time. Anticipation of an event at one point relative to another point, among these six, and appropriate simple action to record the accuracy of this anticipation, constituted the essentials of the test.

Two chain belts, driven at slightly different speeds by a reducing gear from a small motor, moved around the test area. A red flag was mounted on the belt moving at the higher rate of speed, while a white flag was mounted on the other. Electric contacts affecting each belt when the flag was in their two-inch area were placed at the six points indicated by white strips in the illustration. When both flags were on the same white area, or when the two flags were on any two white areas at the same moment, the light in the center of the test would flash.

Six switches were mounted in the top of the test, each controlling the contact of the belt carrying the red flag in the corresponding location around the test. If the testee pulled the appropriate switch when a contact was about to be made, the light would not flash. This would be a successful performance. If he failed to pull the switch, pulled the wrong switch, or pulled the right switch too late, he scored an error.

A cycle of operation lasted 2 minutes and 57 seconds, from the time the red flag overtook the white flag until the red flag overtook the latter again. During a cycle the light would flash nineteen times if no switches were pulled. Scoring was therefore simple. A perfect score would be just nineteen pulls, with no lights flashing. Pulls in excess of nineteen would be unnecessary and would thus count against the testee. Similarly, lights allowed to flash were failures and counted against the testee. The record was taken of the total number of pulls per cycle and of the number of lights flashing per cycle.

Some persons observing the operation of this test have inferred, incorrectly, that the testee was required to learn from experience during the testing which switches controlled the light contacts. This is not a learning test. Specific directions on operation were given, accompanied by supervised practice during one cycle. If this was not sufficient to develop the necessary insight—a condition met very rarely—a second cycle under supervision was given.

During the first experimental applications of the test to some 60 operators and other persons, three complete cycles were given and recorded. This number was later reduced to two, due to the slight improvement found in successive trials and to the high measured reliability of the test. Also, lights flashing were weighted two points each, as against one point for each excess pull. This seemed to be a reasonable valuation. A high numerical score, therefore, represented many excess pulls, many lights, or both.

On the three successive applications, for 69 operators for whom records were available, the average number of lights was 8.57, 7.35, and 6.30, respectively, with variabilities of 3.59, 3.33, and 3.67, respectively. The average number of excess pulls was 12.5, 11.34, and 10.71, respectively, with variabilities of 7.00, 7.14, and 6.09, respectively. The small changes from trial to

trial indicate the high reliability of the test and the absence of significant gains through prolonging the testing time. Intercorrelations of numbers of lights with numbers of pulls were .46, .49, and .48, respectively, on the three trials. Intercorrelations of light scores were .61, .54, and .53, and of pulls .53, .65, and .67, respectively. Total scores gave .87, .82, and .79, respectively (first with second, first with third, and second with third).

TEST RESULTS

From the outset of experimentation with this test, the results were so unique and promising that an extensive study was made of the scores in relation to other tests applied to the same operators. The results are summarized in Table IX, above. This was not merely a research venture but a program to accumulate data about each man, by means of which the factory manager would later choose, from among the present operators, the men to be transferred to a new plant, and by which he would select new operators from among the more than 125 applicants tested during the summer and fall of 1940.

TABLE IX.—MACHINE-OPERATOR TEST SCORE CORRELATIONS WITH OTHER MEASURES

Other measures	Foreman's ratings	Machine-operator test scores
Foreman's ratings		.40
Pin Board time	−.02	.44
Pin Board complexity	.06	.23
Visual perception (average time and error percentiles)	.35	.43
Intelligence	−.12	.26
Machine-operator (average two trials)	.40	
Detroit Mechanical Aptitude	−.02	.48
Machine-operator improvement	−.33	.03

The significant correlation of scores on this test with scores on the Detroit Mechanical Aptitude Test is of considerable interest. Here is a nonverbal performance test of distinctly dynamic character, involving perceptions of space and time relationships and manual or motor responses, correlating .48 with a verbal, or

largely verbal, test of "mechanical aptitude." If it can be shown that this test, in less than 10 minutes of employee testing time, can effect a satisfactory classification of personnel with respect to this ability, we have the basis for the development of an infinite variety of tests of this general type for the selection of personnel for similar jobs in industry.

It must be noted that this test does not measure all the specialized abilities called for by the job. These operators are required to maintain a frequent inspection of the material in the process of manufacture. This involves general visual and kinesthetic perception of the sort required by the inspector-packers at the end of the machine. The operators must perceive material and product running off-standard and must make the appropriate adjustments or shut down the machine to have such adjustments made.

The operators must be physically able to handle the large and heavy rolls of material fed into the machine. They must also be active in moving along their machines to maintain their inspection of the product in process of manufacture and to view the rolls of material that are becoming exhausted. Ability to remember which rolls will run out first, and perception of time during the intervening period, are important and are also included in the requirements on the test.

While little confidence was placed in the foreman's ratings as a criterion by which to evaluate the success of the test, a comparison of such ranks or ratings against relative ranks on the test is of some interest. Figure 12, page 72, shows this relationship for 34 cases in the early stages of this testing. From this it is apparent that, if a testee is in the upper half of the tested group, his chances of being in the upper half on the foreman's ratings are 15 out of 17. If he is in the lower half of the tested group, his chances of being in the upper half of the foreman's ratings are only 6 out of 17. A high ranking on the test is thus good assurance of high ranking by the foreman.

For 45 employees for whom data were available, the intercorrelations of scores on the Machine-operator Test with other measures are shown in Table IX, page 70. It is quite significant that the figure is .43 for the combined time and error scores on the Visual Perception Tests and only .26 with scores on the intelligence test—in this case Wonderlic's adaptation of the

Otis test. This tends to support our inference that performance on the Machine-operator Test depends upon the process of perception rather than upon reasoning and closely related abilities. On the other hand, the low correlation with Pin Board complexity scores indicates that speed in complex hand movements is not very important to success on the test.

A comparison of the measures of success on the test with the same measures of success on the job is also of interest. Thus, this test correlates with foreman's ratings .40, which is not

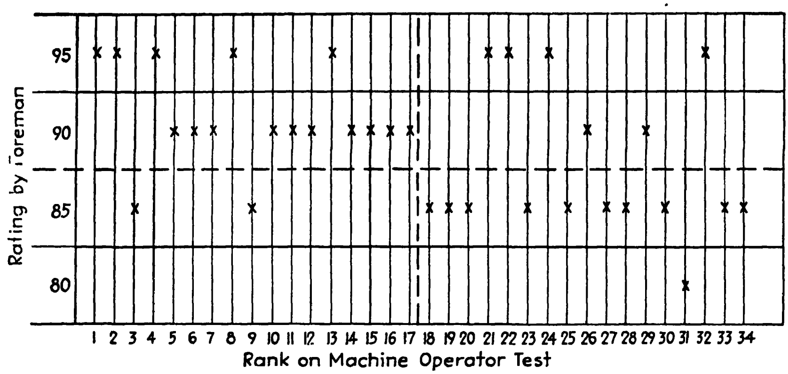

FIG. 12.—Comparison of foreman's ratings with ranks on Machine-operator Test.

impressive. In spite of this low figure, a study of the individuals at the high end of the classification indicates that the test is doing very well what it is intended to do, that is, to select the applicants who will have a high assurance of success on the job.

A study of the correlations of the several tests with the foreman's ratings is of some interest. Thus, neither Pin Board complexity scores nor the Detroit Mechanical Aptitude Test scores are significantly related to these ratings. Intelligence test scores show a slight inverse relationship with the foreman's ratings, as do the Pin Board time scores. The Machine-operator Test improvement scores, which reflect the ability to achieve a better score in the second trial as against the first trial in this test, are also inversely related to the foreman's ratings. Average scores on the Visual Perception Tests correlate .35 with such ratings, comparing favorably with the .43 correlation between

these tests and the Machine-operator Test scores and tending to confirm the hypothesis upon which the latter test was designed—that the most essential element in the job is visual perceptual ability.

The foregoing considerations serve to emphasize still further the necessity for the separation of perceptual factors from manual factors in the analysis of jobs for the purpose of test construction. It is quite apparent that manual dexterity is of some importance as a factor in success on the test, but visual perceptual ability is important on both the test and the job, and constitutes the only significant point of agreement among these results.

If one were using the older technique of weighting tests by the regression method against the foreman's ratings as a criterion, he would have left only three measures for selection purposes from among the seven shown in the table. Of these three he would be compelled to value the visual perception measures quite highly, and could be forced to the wholly illogical equal weighting of the improvement ratio scores on the Machine-operator Test, while he weighted only lightly the ordinary scores on the Machine-operator Test.

Yet the graphic analysis shows very clearly that, if we take for employment only the high-ranking applicants on this Machine-operator Test, we shall get a very high proportion of successful machine operators. If we add to this the requirement of high average scores on the Visual Perception Tests and high scores on the Pin Board Test, we shall achieve an excellent selection from among the group of applicants.

It is a rather curious fact that the addition of the scores on the Detroit Mechanical Aptitude Test will contribute little to the excellence of the selection obtained by the foregoing three tests, yet the total time for these three measures is considerably less than that required for the administration of the Detroit Mechanical Aptitude Test.

It is by no means certain that the Machine-operator Test includes all the elements that we should like to test on the job. Whereas this test is highly reliable as a measure of whatever it does measure, there must be other factors conditioning success on the job not reflected in the performance called for on the test. Certainly there is nothing to measure the employee's

ability to move about, to resist fatigue, to adjust to the social situation of the work place, and many other factors that one might analyze out of the job itself. Each of these, and perhaps many others, must be of some importance, but we are thus far without adequate measures of their effects. Whereas some of them could be added to the test or measured by separate tests, it must be remembered that the law of diminishing returns operates here as elsewhere in human affairs, and the practical job of employee selection will justify only a certain degree of complexity, a certain amount of expense, and a limited period of time.

Certainly this test has justified itself as an instrument for selection on the grounds that it is quick, accurate, and economical. In addition, it has opened the way to a whole new field of industrial testing which may now be further explored by those who care to make use of this promising technique.

CHAPTER VIII

SOME STANDARD JOB TESTS

The following brief descriptions of a number of tests designed and applied during the past several years may be suggestive. They are not presented as examples to be copied and applied in different situations, but rather to indicate the variety of designs that may be evolved to solve particular industrial test problems.

Sewing-machine Test

The test for selection of sewing-machine operators is shown in Fig. 13, below. This was designed hastily for the selection

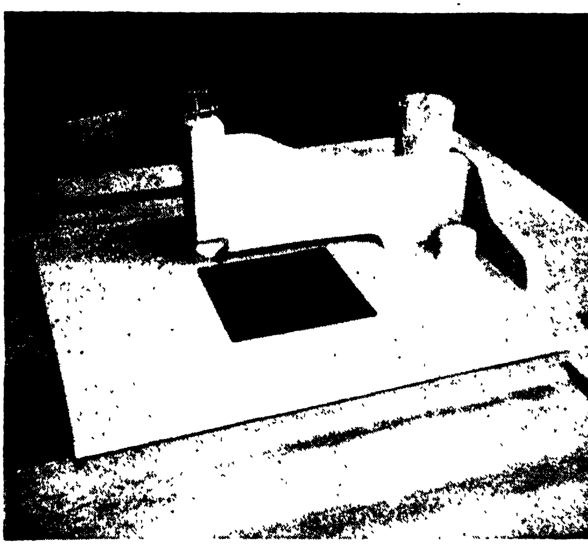

Fig. 13.—Sewing-machine Test.

of operators and trainees for sewing operations on gas masks. The material to be handled was coarse and stiff. The operation placed the burden of manipulation upon the left hand, which is

also true of the majority of operations on the standard power sewing machine. The test was accordingly designed to require the larger part of the necessary manipulations to fall upon the left hand, by placing a fence rather low on the projecting arm. Two different mazes were made, one with straight paths and right-angle turns, the other with curved paths and acute-angle turns.

Clumsiness in running the mazes through this test often results in attempts to make the turns before the ends of the paths are reached. Each time the maze is not strictly followed, a buzzer sounds. This constitutes an error, which may be recorded separately or may be given a penalty value in seconds and added to the total time score.

Whereas this test is still in an incomplete form, it has already demonstrated that its design is basically sound and merely requires extension and refinement to fit it for selection of operators for other types of operations on the numerous varieties of sewing machines now in use.

In its present stage of development, this test does not give adequate attention to the perceptual element in sewing-machine operation. The perceptions required while the machine and material are in motion at various speeds will have to be measured by a separate device which is now under experimental development. Visual Perception Tests A and B afford the best existing measures of this ability.

Motor Rhythm Test

Figure 14, page 77, shows our modification of a test developed many years ago by Henry C. Link. It consists of a phonograph turntable rotating at 40 turns per minute and 50 half-inch steel balls. Operation involves attempts to drop the balls one at a time through a slot in the rotating disk. Balls dropping through the slot fall into a compartment near the testee. Balls failing to drop through this slot roll into another compartment on the side of the test. Scores are in terms of the percentage of success attained.

It is clear from Table II, page 27, that scores on this test are not significantly related to scores on any of the other tests. Yet the test has fair reliability and could be made to show quite

high reliability by prolonging the practice period before testing and by increasing the number of balls dropped for scoring. Simplifying the operation of dropping the balls would also increase the reliability of the measure.

Although this is not the best test that can be designed for measuring the rhythmic function that seems important in cer-

FIG. 14.—Motor Rhythm Test.

tain operations, it has proved useful in selection. The failure of its scores to correlate significantly with the scores on other tests is of great theoretical and practical importance in industrial testing and deserves further study.

HAND-FOOT TEST

The test shown in Fig. 15, page 78, was designed primarily to select operators for simple work on kick presses. It therefore has a very short hand cycle with a simple foot element,

differing markedly from the cycles of the somewhat similar tests shown in Figs. 1, 2*a*, and 2*b*, pages 28, 30, and 31. This test affords an excellent measure of hand-foot coordination ability free from the bilateral and dual elements that complicate the other hand-foot tests.

FIG. 15.—Hand-foot Test II.

HAND DEXTERITY TESTS

Pin Board I, shown in Fig. 16, page 79, was designed to select trainees for an operation requiring one and two one-hand elements, intervening between two bilateral elements in the cycle. The first board features the basic two-hand elements uncomplicated by the one-hand element. The second board introduces one one-hand element by requiring the testee to

place one medium-sized pin in a hole after each pair of large-sized pins has been placed. The third board requires the plac-

FIG. 16.—Pin Board I.

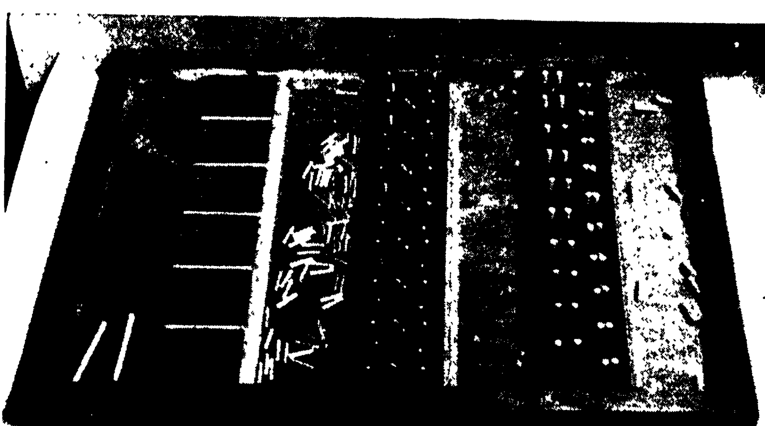

FIG. 17.—Pin Board II.

ing of a medium-sized pin with the right hand, followed by a small pin with the right hand also, between the large pins which

are placed two at a time as before. The left hand remains at rest during these operations.

Two kinds of scores are obtainable from this test. The first of these, the over-all time for the three boards, is the more significant for most jobs. The second, derived from the increase of time required by the third board over the first, constitutes a set of scores only moderately related to the total time scores. These latter scores reflect the dexterity of the testee on increasingly complex tasks involving one-hand elements in the operation cycle and are valuable indexes of inherent individual differences.

A similar test, shown in Fig. 17, page 79, was designed for certain assembly operations in gas-mask manufacture. The ability to work with both hands at considerable speed, and particularly to keep one hand from interfering with the other, seemed most important on this job. The testee is permitted the usual practice before testing and is then allowed to pick up with each hand as many pieces as he can manage at one time. Only the total time score is considered here, *i.e.*, the sum of the separate times on the three boards.

Inspection Test

The Inspection Test, shown in Fig. 18, page 81, was designed for measuring the inspection ability called for on the same job for which the test shown in Fig. 16, page 79, was designed. This ability is both visual and kinesthetic.

The 100 boards constituting the test are divided thus: 60 standard, 20 shorter than standard by various small amounts, and 20 longer than standard by similar amounts. The testee is required to sort these boards into three piles according to size, making his perceptions quickly, since he is scored on both time and errors. Error scores are quickly computed by turning each pile on its side and spreading the boards to expose the backs, which are marked in distinctive colors for the several sizes.

A test of this sort, numbering only 100 items, is usually too short to have high retest reliability. This can be corrected by giving the test a second or third time, or, of course, by increasing the number of items. Usually, however, it is found that

high scores on the shorter test are made by the same persons who make high scores on the longer test. Since we are seeking only the best, lengthening the test is seldom necessary except for research purposes.

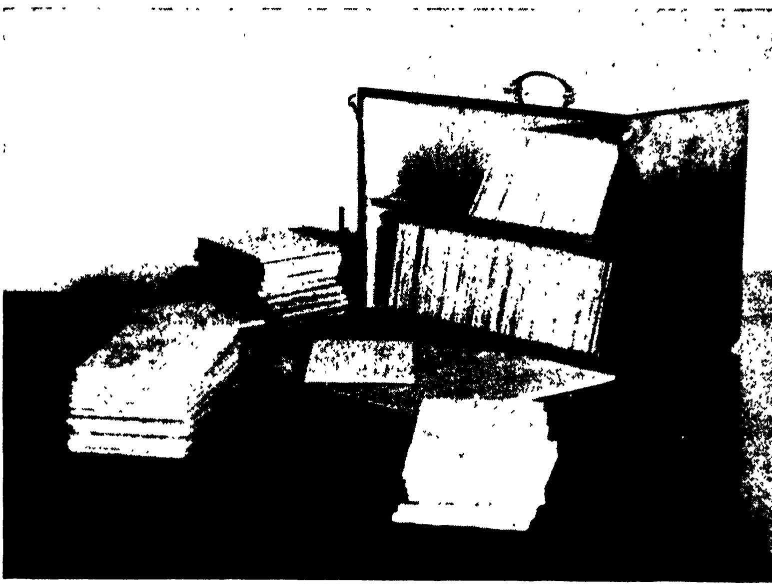

FIG. 18.—Inspection Test.

DUAL HAND TEST

The simple short-cycle Dual Hand Test, shown in Fig. 19, page 82, has practical utility in selecting dual operators for short-cycle work. It is also most useful in demonstrating the savings that can be effected by dual operation over one-hand operation in the simplest cycle. Its operation involves merely picking up two marbles, one with each hand, from the rear compartments and dropping them through the two rear holes; then, similarly, picking up two marbles from the front compartments and dropping them through the two front holes.

Four compartments are used rather than only two to introduce the attention shift required on the job for which this test was originally designed. It can, of course, be performed using only two compartments. The marbles are of four colors, mak-

ing separation and assessment of penalties easy in case some are placed in the wrong holes. Only a total time score, after practice, is taken. The test is quickly reset by raising the central container about one inch, whereupon the marbles roll out into their several compartments.

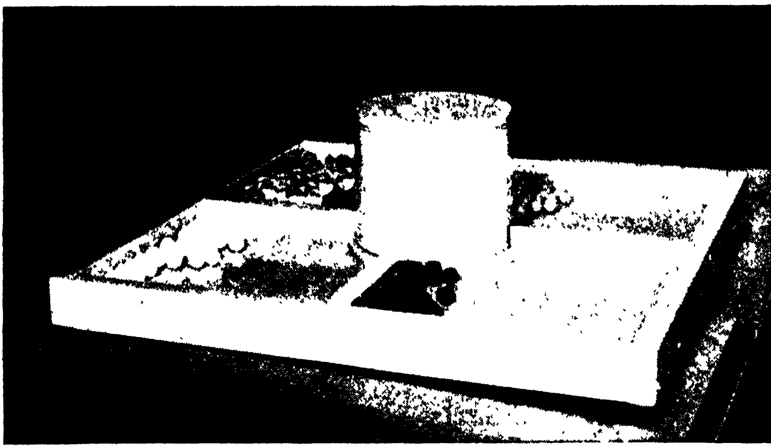

FIG. 19.—Dual Hand Test.

BILATERAL HAND TEST

The test shown in Fig. 20, page 83, was designed for the selection of operators who were required to pack a variety of different items in the same package. These operators worked with both hands, taking three items with one hand and one with the other, or two with one hand and three with the other, etc., and positioning them in the boxes.

Items in the hoppers, four on each side, are placed several at a time in four sets of holes in the board in the center of the work area. When filled, the board is moved laterally, dropping the items into similar hoppers below. The lower hoppers may be interchanged with those above at the end of the test, thus making the test ready for the next testee. Items are of different colors, sizes, and shapes.

The Bilateral Hand Test affords a total time score which does not correlate very highly with the scores on other tests. The test, however, permits a qualitative analysis of particular

disabilities in grasping and positioning as affected by the necessity of handling several items at one time. This latter is often highly significant. Good bilateral operators seem to be as scarce as good dual operators. Yet the really good operators in both

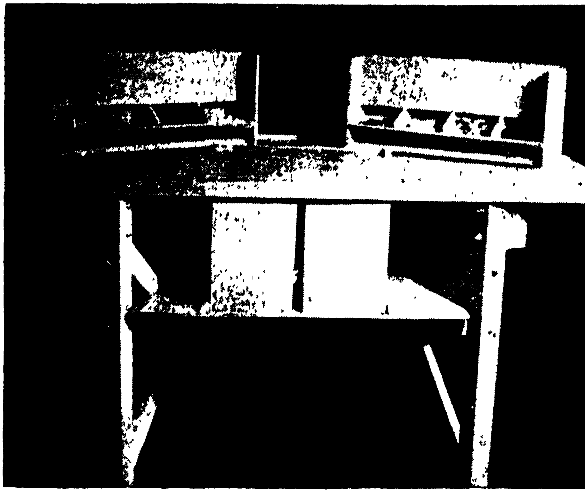

FIG. 20.—Bilateral Hand Test.

categories perform the tests for the respective abilities with ease and amazing rapidity.

CONCLUSIONS

The tests shown in Figs. 16 and 18, pages 79 and 81, were the first ones designed strictly from the operation analyses made by methods engineers. They were highly successful in identifying the best operators then on the jobs, thus helping to convince the management of the company of the efficacy of this technique. When used in the selection of new operators, they were similarly successful.

Following the success of this early testing, the procedures were extended to include the new plants then being erected, both to select operators for transfer from the old plants to the new and to select new operators for training. New tests were designed for types of jobs that occurred in sufficient numbers to warrant the expense.

It has been demonstrated beyond question that a combination of improved methods and selection of operators by appropriate tests will effect large increases in production. Just how large the increases have been in particular situations it is not ethical to state. Furthermore, the selection process alone has resulted in substantial increases in production where no changes in methods were made.

In view of these facts, there may be a strong temptation to imitate the tests herein described and to apply them in other situations. This is almost certain to lead to disappointment. Unless the new situation is parallel with the one for which a test or tests were designed, the selection process may not be valid for the jobs.

We have steadfastly refrained from giving specifications and norms for the foregoing tests to anyone. The tests belong to the firms for which they were designed, and the norms are matters of internal interest only. It would be far simpler for a competent person to design new tests than to try to adapt these tests and norms to a different situation.

It must be remembered that, if the basic design of a test is sound, any desired reliability can be obtained by increasing the length of the measure. The length chosen will always represent a compromise between high reliability and the reasonableness of the testing time. Too much time must not be demanded, or the testing procedure defeats its own end.

The best measure of validity of a test for practical purposes is the proportion of the selected group—the highest 50 per cent, 25 per cent, or even 5 per cent—who make good on the job. The usual correlation measures take into account equally the persons at the lower end of the ability scale, individuals in whom we are not interested because we have rejected them.

The technique of test design has now arrived at the point where we no longer ask of a measure, "Will it work?" The question now is, "How well will it work?" If the design is right the test must work. A test related by basic design to the job for which it is intended must give a significant distribution of talent with respect to the basic abilities called for in performance on the job.

CHAPTER IX

MOTOR SKILLS AND PRODUCTION

It was pointed out in Chap. I that the usual criteria of success in industrial employment are wholly inadequate as standards against which to measure the effectiveness of aptitude tests. It was argued that appropriate tests yield results that can easily and reliably be used as the criterion against which the effectiveness of management, especially the quality of supervision, can be evaluated. This is an unorthodox proposal and deserves further examination.

Any work-limit type of test, involving a certain number of items to be manipulated and scored in units of time, is subject to having its raw scores interpreted in a variety of ways. We may simply and arbitrarily select a time score below which no applicants will be accepted and above which applicants will be accepted if they pass certain other tests and the personal interview. The establishment of such a critical score on the test will be determined by a variety of factors. If applicants are numerous and the jobs to be filled are few, the standard will be set quite high. When many jobs are to be filled and the number of applicants is small, the critical score is likely to be moved downward.

A practical safeguard against this situation is to build up a file of qualified applicants and draw them from this file in order of excellence to fill the labor requisitions. Another safeguard is to make sure, through suitable publicity, that applicants will be numerous enough to yield the number of highly qualified individuals required.

The arbitrarily determined critical score is a useful rough-and-ready technique for employee selection, but used alone it tells us little about the applicant who is much above or much below the designated point. Two other devices are in common use to give additional information. The simpler of these is the percentile score, expressing the percentage of the tested

group falling below any particular score in such group. Thus, if a testee scores 86 on the percentile scale, this means that 86 per cent of the group to which this individual belongs scored below him and only 14 per cent scored above him. It is understood that where time scores are used the shorter times are the better scores.

Standard Scores

The other technique involves changing all raw scores into sigma scores with an average of, say, 50 and a standard deviation of 10. This conversion process is simple, but the interpretation of scores is not simple. It possesses, however, a marked advantage over percentile scores in that it makes a unit of the standard score equal to every other unit of score along the scale. Thus, the 3 units of ability between scores of 37 and 40 on such a scale are equal to any other 3 units, such as those between 73 and 76. This is not true of percentile scores, which fact is the most serious objection to their use.

Both the foregoing techniques permit direct comparison to be made between the individual's performances on several tests. A percentile score of 76 on one test means the same as a percentile score of 76 on another test. The same is true of equal standard scores. Both permit graphical representation of an individual's ability as measured by several tests with respect to some common average.

Both of these methods seem quite simple, obvious, and easily interpreted to the professional psychologist or the statistician, but they are not familiar to most engineers and employment personnel. We have, therefore, decided to abandon both and to adopt a new method of raw-score conversion in our work.

Percentage Efficiency Scoring

It is observed that, in tests of the work-limit type, the mean, or average, score can be used as the 100 per cent level. Scores above and below this average can be converted into terms of percentage efficiency merely by dividing the raw score into the average score. The resulting quotient when multiplied by 100 is the percentage efficiency figure that stands for the raw score on the test. *This figure is easily computed, always means*

the same thing, permits direct comparison of the test scores, and has the additional advantage of being thoroughly familiar to most of the people who will use it. Table X, below, shows the percentage efficiency technique applied to a distribution of scores.

Effects of Length of Cycles

On motor tests we have seldom found such percentage scores to range higher than 200 per cent. The lowest individual percentage score thus far obtained is 42 per cent. It is observed

Table X.—Distribution of Scores on Motor Rhythm Test

Number scored *	Frequencies	Percentage efficiency
38–39	2	157
36–37	2	149
34–35	8	141
32–33	11	132
30–31	5	124
28–29	6	116
26–27	13	108
24–25	15	100
22–23	14	92
20–21	13	84
18–19	9	75
16–17	7	67
14–15	3	59
12–13	5	51
10–11	1	43
	114	

* Average = 24.55

that the number of percentage points in the range tends to be inversely related to the number of elements in the test cycle. That is, the greater the number of elements in the cycle, the smaller the range of scores. This is shown in Table XI, page 88.

The practical application of this latter fact is that we will find the greatest differences in ability among employees working on simple jobs and the smallest differences among employees on jobs where the operation cycle is long and complicated.

Nevertheless, we must not overlook the fact that the range on the most complex cycle is from an efficiency of 62 per cent in the poorest person to 155 per cent in the best person. While relatively greater gains in production will result from selection by test of the employees on the more simple jobs, the gains on the more complex jobs are only slightly less spectacular. The figures in Table XI, below, indicate that for the three most complicated tests the maximum production is 55 per cent,

TABLE XI.—RANGE IN RELATION TO NUMBER OF ELEMENTS IN CYCLE

Number of tests	Number of elements in test cycle	Minimum and maximum percentage scores	Range in percentage score points on each test
1	9	62–155	93
1	7	76–167	91
1	6	80–171	91
3	4	42–163 54–186 50–185	121 132 135
1	3	53–212	165

67 per cent, and 71 per cent above normal, while for the four least complicated tests the maximum production is 63 per cent, 86 per cent, 85 per cent, and 112 per cent above normal.

It must be emphasized that these superior performances indicate potentialities, and not actualities on the job. The figures result from performances under the most favorable incentive conditions and during a comparatively short period of time. It may be inferred, however, that the optimum rate of performance on the job for each individual will be roughly proportional to his percentage efficiency figure on the test, *i.e.*, with suitable incentives and under competent supervision.

INCREASING EFFICIENCY

It is possible, both theoretically and practically, to build up a work organization of personnel none of whom is below a poten-

tial percentage efficiency of 130 per cent. From such an organization a minimum net gain in production of 30 per cent could be—and should be—obtained.

The foregoing observations carry still another implication for production. If the shorter cycles show greater ranges, and thus greater potential gains on the part of the best persons tested, the possibility of further gains in production by merely breaking down the more complex cycles and having the work performed in two or more shorter cycles, usually by two or more operators, is evident. It would thus seem that gains in production through division of labor are in part due to this characteristic of the shorter cycle.

As a practical testing problem it is evident that it is easier to design and construct tests for the shorter cycles. It is also possible to find a much larger proportion of qualified applicants for these cycles. Training time can be reduced, and a high level of production can be more quickly obtained.

It is also evident that an operator of suitable dexterity can learn a number of short cycles in the time he might otherwise devote to the mastery of one longer cycle. The shorter cycles, therefore, make for more flexibility in the labor force. Operators can be switched from job to job with little adverse effect on production.

Under certain conditions it may be equally important to break down the longer cycles to avoid too great dependence on key operators on important long-cycle jobs. This is especially important where the operations are performed on or along conveyors. We have seen plants in which a few highly skilled operators on long-cycle jobs could easily cripple or completely tie up production, if they wished to do so.

The two most certain safeguards against the foregoing conditions are to train understudies in considerable numbers for such jobs, or, better still, to break down the long-cycle jobs into several short-cycle jobs. The only plausible argument against the latter procedure is that it tends to make the jobs more monotonous and, inferentially, to lead to undue fatigue and higher labor turnover.

Monotony

The increased-monotony argument does not seem, from our observations, to be valid for properly selected operators. We have seldom found operators who preferred the longer cycles, except for reasons of more pay or greater social prestige among their fellows. The longer cycles are demonstrably more fatiguing and impose strains upon attention and perception such that relatively fewer operators are able to tolerate them.

The real cause of maladjustment on a repetitious job seems to reside in two factors. Either the operator is too intelligent for the job, or the pace of the work is beyond the worker's optimum. In the first instance the error of selection can be easily corrected by transfer, and proper safeguards against such situations can easily be set up. In the second instance the only safeguard is the prior determination of the individual's ability by suitable tests and his assignment in accordance with these results.

Repetition and Fatigue

There has been in the recent past a considerable amount of sentimentalized discussion of the fatiguing effects of repetitious work. Much of this seems to have come from persons who were themselves unfitted for such work, owing to one or the other of the foregoing factors. Introspection and reflection upon their own brief and unhappy experiences have led them to the error of generalization from too few cases. Their arguments have had, nevertheless, a certain amount of plausibility for other persons of mental organization similiar to their own.

From some of the more academic discussions, one may conclude that most factory work is highly repetitious and therefore open to condemnation for implied social effects. The fact is, however, that only a small proportion of jobs in general is of the short-cycle, repetitious variety. It is also a fact that it is very easy to find the proportion of qualified operators who actually prefer such jobs.

It is likewise important to remember that a job is not an end in itself—except for a few fanatics—but is a means to other

more desirable ends. These real ends justify and make tolerable the means that may seem onerous to the few who, without prompting, really think about them. With the shorter work day and the longer avocational and recreational day, the monotony argument loses its plausibility almost completely.

Ability and the Shorter Day

The shorter work day increases the need for more careful and logically defensible techniques of selection. Individuals of better mentality as well as of higher work aptitudes will

Table XII.—Percentages of Specialized Abilities Falling in Highest Quarter of Various Test Results

Tested abilities	Eagle Pencil Co., New York, N. Y., 107 applicants	Johnson & Johnson, New Brunswick, N. J., 121 operators	Johnson & Johnson, New Brunswick, N. J., 130 applicants
Dual	21	16	31
Inspection	25	27	23
Dual and inspection	8	13	12
General dexterity and inspection	20	19	22
Hand-foot and inspection	8	15	12
General dexterity and hand-foot	22	20	21
Dual and hand-foot	9	14	22
General dexterity, hand-foot, and inspection	4	10	12

accept jobs during a six-hour day that they would not accept or continue on during an eight-hour day. Proper selection and placement can easily achieve a daily production for the six-hour day that will exceed the usual production for the longer day. The length of the work day is one of the incentives to consider in formulating a wage policy.

As the work day becomes shorter, more concern is felt about the actual distribution of individual abilities among the operators and applicants for jobs. Heretofore, it has not been known how many among a hundred persons were possessed of the basic aptitudes for inspection, dual operation, hand-foot operation, bilateral operation, or even general finger and hand dexterity. To throw some light on this situation, several analyses were made of groups for whom extensive test results were available.

Table XII, above, shows the percentage of persons in

each group whose specialized abilities or combination of specialized abilities were in the best quarter of their group. It is at once apparent that some of these abilities are more rare than others and that certain combinations of these abilities are less frequently found than other combinations. The futility of the effort to find these rare individuals by interview or ratings on the job should be equally evident.

Reassigning Employees

Since the distribution of these abilities is not materially different for experienced operators than for new applicants, it may be suggested that considerable increases in production could be achieved in the average plant merely by reassigning the present operators to the jobs for which they are best fitted. In one situation familiar to us, it is calculated that a potential increase of 17 per cent in production would have resulted from such a change. This potential increase could have been made actual by suitable incentives and supervision, without any change in methods of work on the existing jobs.

From the standpoint of the foreman, it is an obligation of management to furnish him with operators who meet the requirements, in terms of special dexterities and perceptual abilities, for a high level of performance on the jobs he has to fill. A foreman's primary responsibility is to get out production. If he is given employees of low potential ability who require unduly long and uneconomical training periods, his true function is severely interfered with. He has every right to demand the application of the best available techniques in furnishing him with the human tools in the productive processes for which he is responsible.

Selection of Foremen

This raises a question about the selection of the foreman himself. Under present conditions it would seem that it is far more important that the foreman should know the techniques of supervision and training than that he be an expert on one or more of the jobs under his supervision. The capacity for learning new methods, employing new techniques, and adopting new points of view on policies and practices is far more im-

portant than the sentimentalized craftsmanship of the old-fashioned sort.

Methods of work undergo frequent change. The increasing use of power machinery and the minute subdivision of labor, together with instrumental control of quality through the use of precision instruments, make increasingly important the matter of supervision.

The emphasis is shifted from quality to quantity, since the former is largely removed from the voluntary control of the operator and assumed by the machine. The specifications for operators, therefore, feature the dexterities and motor coordinations which will best serve to meet the demands of quantity production. Such operators can be most easily and quickly trained in the particular methods of work employed on the job. They are the ones who will do this work most easily at the required rate, *i.e.*, with the least stress and strain.

CHAPTER X

TESTS AND LEARNING ABILITY

It was pointed out in Chap. VI that on certain dual operations the individuals selected by test began production on the job at a higher level, gained in skill at a more rapid rate, and finished at a much higher level than those who had made below-passing scores on the tests but were also employed.

The situation confronting us in such training is physiological rather than psychological. Every individual has, for all reactions of his organism, a maximum level of performance called the "physiological limit." Beyond this point, determined by heredity rather than by environmental factors, he cannot go, regardless of the amount and quality of his instruction and coaching. It is reflected very clearly in the maximum record of the high jumper under expert training, and in the best attainment of a runner in the 100-yard dash or the mile run.

Sustained Production

This limit imposed by nature is of similar importance in industry. On any job there is a maximum limit beyond which any particular operator cannot go. The problem then becomes one of selecting the person with the highest measured aptitude for a given job and subjecting him to the training and incentives necessary to attain a high level of performance.

It should be noted that we have not said "to attain his maximum level of performance." Under exceptionally favorable circumstances, and for short periods of time, most persons can reach such a maximum level of production—a level that cannot possibly be maintained for a long period. Thus, we have trained operators to a level that indicated production at 150 gross of the product per day, if that level could be maintained. On sustained production these operators produced 90 gross per day under training supervision. Placed on identical jobs in the

production departments, they soon declined to a level of 50 gross per day, which was the average level of production on those jobs in the plant.

This decline in production was caused by two factors external to the operators. One was coercion of the "don't spoil the rate" sort by fellow workers. The other was pressure by foremen who thought that this higher production reflected upon their own ability to get work out of their older employees and who therefore imposed handicaps of parts, materials, and fixtures (shortages and poor tools), and made remarks calculated to discourage such efforts.

The first factor represents the result of lack of confidence on the part of the workers, an attitude born of previous experience with a management that abolished jobs and cut rates when it thought that the earnings were too high. Almost every job has a commonly accepted maximum, recognized by the employees and enforced by their public opinion. They have learned that persistently exceeding that maximum will lead to reduction of the rate or change of the job setup.

Employees also have a certain self-respect with regard to their work which makes them resent a performance that "shows them up" or makes them look inferior as producers. A shop is a social unit as well as an economic unit, with all the stresses that characterize social units outside the plant. This is a point that should be emphasized in all management and foremanship training.

The second factor, the attitude of foremen, reflects poor organization, poor supervision, and poor morale. There is no remedy for the situation save selection of more competent personnel in the foremanship positions and better management in general. Continuous foremanship training, when coupled with close supervision of foremen, is the best safeguard that can be offered under the usual conditions in industry.

In such situations the pacing of work by the machine offers decided advantages. Assembly on and along conveyors, belt and dial feeding, and packing from the end of the machine are good examples. Operators selected by tests for the dexterity and rate of speed demanded by the machine can work with far greater freedom from the strain that always attends jobs on which the pace is determined by the operator's own volition.

Such strain is not relieved by the hourly posting of production figures, although this procedure does tend to add an incentive and to hold up the level of production on this voluntary basis.

The real problem is primarily one of selection. It is the problem of selecting operators of high basic aptitude and of giving suitable training to attain their optimum skill. The optimum level will, in general, be higher for those of higher measured aptitude. Also, the training task will be easier and shorter.

Effects of Low Aptitude

In the evaluation of the attempts to train operators of low measured aptitude to a high level of skill, two conclusions stand out. The first is that such operators learn more slowly and with much greater difficulty than those of higher aptitude. Often such learning is achieved at a cost so great, in time, spoilage, supervision, and instruction that it is an uneconomical proposition. The other conclusion is that operators of low measured aptitudes seldom or never attain the level of those of higher aptitudes. Where they do succeed in attaining a good level of performance, it is usually maintained only temporarily and with excessive effort.

It thus seems impractical and unwise to expend much training effort on individuals of low aptitude. It is far better to assign such persons to simple jobs more suited to their limited abilities, or else to reject them in the first place. It is no kindness to anyone, at any level of ability, to assign him to a task beyond his level, no matter how desirable he may otherwise be or how great is his need of work.

Still worse is the procedure, formerly more common than now, of trying out a number of individuals, and dropping the many who cannot make good within a short time. We have definite statements of employers that only one among three accepted applicants makes good within six weeks on certain sewing-machine operations. We have other evidence concerning jobs on which only one person in five or six can, after a period of training and experience, actually attain the prevailing standard. Such a mortality rate is neither sensible nor humane.

Results of suitable tests, when properly interpreted, give at least four correct selections out of five, and even nine out of

ten. The ones who fail after selection by test and interview do so not from lack of the particular abilities measured, but usually for personality reasons not caught in the interview. *The test is a positive guarantee of the presence of the ability at a level that may be determined. Whether that ability will be developed depends upon training and supervision. Whether it will be effective in production depends upon management, particularly upon supervision.*

The Average as Ideal

Probably one of the greatest drawbacks to production increases is the willingness on the part of management to be satisfied with average or mediocre performance. When this mediocrity becomes an ideal to be achieved and maintained, as it often does, there is little hope of attaining those higher levels of production that the innate abilities of some of the personnel warrant.

It is clearly demonstrated in both the test results and in the actual increases in production accomplished by selection based on such test results that potential gains are sufficient to offset almost any other competitive disadvantages from which an organization may suffer. We once received a letter from an individual who disapproved of selection by test, on sentimental and ethical grounds. He destroyed the logic of whatever case he made by citing, at the close of his letter, the instance of a company well known to him that "was unfortunately enjoying great prosperity in a highly competitive field" and that was making use of a rigorous test-selection technique.

We once inquired pointedly of the production manager of a plant we had visited, "How can you maintain your competitive position in the industry with so many obsolete machines and so many old-fashioned methods?"—a fact he had previously pointed out, with illustrations. He replied, "We can get away with it because our competitors are no better and usually are somewhat worse." A subsequent visit to one competing plant abundantly confirmed his conclusion as it applied to that competitor.

Other Handicaps

Some plants are severely handicapped by the poor quality of the applicants they are able to attract to their employ. We were much perplexed by this situation in the early days of the attempt to build up the personnel of the gas-mask plant in Chicago. The difficulties of transportation to the plant site certainly constituted one reason. Another was the lack of adequate publicity on the kind of labor required. Still another possibility that could not be readily confirmed was the fact that many potentially satisfactory applicants might fear work in a plant having some connection with poison gas.

Fear of the work or of certain machines must not be overlooked in selection as well as in the job setups. We have observed that employees resisted assignment to certain jobs both directly and indirectly for this reason. Some have told us flatly that they did not want to learn to run certain machines, particularly certain paper-box-making machines. We have also noted marked decreases in production on single machines and on groups of similar machines for a long time after an accident occurred to an operator.

Fear and Accidents

Fear of the machine certainly retards learning and predisposes the learner to accidents. These facts are generally accepted. The solution must lie in refusing to assign employees to jobs when such fear is expressed or can be inferred from their performance. Delay in learning, undue fatigue resulting from the strain, and forthright nervousness are signs that should not be overlooked.

Where the test results are analyzed in the manner suggested in Chap. XII for evidence of accident-proneness in the applicant, we have a safeguard against improper assignment. Suitable levels of performance on appropriate selection tests will of themselves effect a reduction of 50 per cent or more in accidents. The additional selection by the accident-proneness technique will accomplish a further reduction up to 20 per cent, or a total decrease of 70 per cent. This result is wholly independent

of other safety measures, including training, as the data were collected in a situation characterized by rigorous accident-reduction measures of the conventional sort advocated by the National Safety Council and administered by a competent and exceedingly painstaking safety engineer.

There is every reason to believe that employees having an accident-free test profile will have the least fear of the machines for which their test scores show them to be competent. We know that they learn more quickly, a result due in part to freedom from the inhibiting effects of fear. It is obvious that this factor of fear must be taken into account at several points in our personnel procedures.

Still another inhibition to success in training lies in previous failures. An employee who has previously failed to meet standards on a job, or on similar jobs, has often built up an attitude of failure. He expects to fail, or fears he will fail, thereby becoming predisposed to failure in this learning attempt. This again reflects the need for selection by test, since the individual should not have been assigned to a job beyond his abilities in the first place.

Inhibitions to Skills

Several other observations from experimental and practical learning situations must be borne in mind. It is often observed that an individual's coordinations tend to break down when he is in a state of rage or of fear. Both states may be brought about by situations in the work place. Tactless supervision, unjust demands, and other similar causes may produce such states. Even the pace of the work itself may impose such a strain upon the person that he becomes emotionally upset to a degree that interferes seriously with his coordinations and therefore with his learning process.

It is also observed that in most learning, especially that involving muscular skills, there occur from time to time *plateaus,* which represent periods during which no improvement is made. These plateaus are well known from intensive studies in the learning of telegraphy, typing, and other similar skills. Escape from the plateaus and resumption of learning often comes through a sudden mastery of a new coordination. This raises

an interesting line of speculation which has direct bearings both upon test techniques and upon time- and motion-study techniques.

In our earlier studies of test results, we found that two individuals might show almost identical speeds in the same elements of the cycle, and that the only significant difference between them was observed when the over-all length of the cycle was considered. It seemed that the important factor, therefore, was not the speed of single elements, which represent simple reactions, but rather the serialization of the simple elements into the complete cycle and the still larger serialization of cycles into a rhythmic pattern.

Effects of Brain Structure

There are sound physiological data to support this conclusion. We have pointed out previously the fact that visual perception is a more complex process than visual sensation. The former takes place in the visual perceptual areas of the occipital lobe of the brain, whereas the latter more simple reaction takes place in the visual sensory areas adjacent to the visual perceptual areas. The situation with respect to complex motor reactions is quite similar. Skilled motor acts seem to be controlled largely by the supermotor area of the brain, which lies in front of the motor area and nearer the higher coordination centers of the cerebrum. It appears, therefore, that the elementary motions or elements of the cycle may be functions of the motor area, whereas the serialization of elements and cycles must be a function of the supermotor area.

This serves to emphasize even further the validity of our conclusion that measured aptitudes are really innate and that they depend upon structures of the brain itself. Modification of these structures is possible only within narrow limits. Lacking the basic structure required, the function will not be performed.

Although it is probably true that an individual has capacity for learning an infinitely large number of skills, and also that certain transfers of skills may take place, it is equally true that each individual is predisposed more to certain skills than to others. Also the limits of his development are implicit and

fixed in the structure of his brain. How far he will develop these innate potentialities will depend upon his experience and training.

That any individual could in a single lifetime exhaust his capacity for developing new skills is wholly unthinkable. Whereas the difficulty of developing new skills seems to increase with age, the general pattern of his set of aptitudes does not seem to change materially with the passing of the years or with the acquisition of numerous new skills.

The extent of the transfer of skill from one job to another is a matter that needs further investigation. There is also a strong suspicion that what may have appeared to be transferred is nothing more than an expression of the same underlying aptitude. Thus, what may seem to be an obvious transfer of skill from fine hand sewing to the handling of surgical ligatures is simply a finger-and-hand dexterity that could be expressed in many similar jobs and is measurable by suitable dexterity tests.

In all situations that call for training in motor and perceptual skills, the best guarantee of success lies in the possession of the underlying aptitude. Evidence of the presence and amount of these aptitudes can be obtained most quickly and economically through the use of suitable tests.

CHAPTER XI

APTITUDE AND SKILL

It has been previously pointed out that the individual of high aptitude starts at a higher level of performance and learns at a more rapid rate on the job than the individual of low aptitude. Differences in job performances between an individual of high aptitude and one of low aptitude may not be great at the start of the period of training, but these differences become more pronounced as training proceeds. The final level of performance of the former will be much above the maximum level that can be obtained by the individual less well endowed.

Many studies have been made of the characteristics of the curves that represent graphically the increase in proficiency in the performance of tasks involving muscular skills. A typical curve of growth in skill for the individual of high aptitude is marked by a rapid rise, with only a gradual falling off in steepness. The typical curve for the individual of low aptitude shows only a gradual rise, with a comparatively early and rapid falling off. The average individual shows characteristics between these two extremes.

The following study of the learning of a muscular skill clearly demonstrates the foregoing. This study is reported at length because the skills developed, while not identical with those common in industry, correspond closely enough to many of the eye-hand coordination skills to make the results representative. This study has the additional advantage of both strong competitive motivation and complete freedom from some of the inhibiting effects of the ordinary industrial position. Moreover, the extent of the skills was measurable in a more exact and objective way than similar skills are measurable on the job. The three individuals taking part in this controlled experiment practiced under standardized conditions. The nature of the apparatus permitted the keeping of numerical scores from which learn-

ing curves could be plotted and by means of which comparisons could be made among the three at various stages in their progress.

Apparatus Used

The apparatus, Fig. 21, below, was a game well known to patrons of amusement parks. It consisted of three round

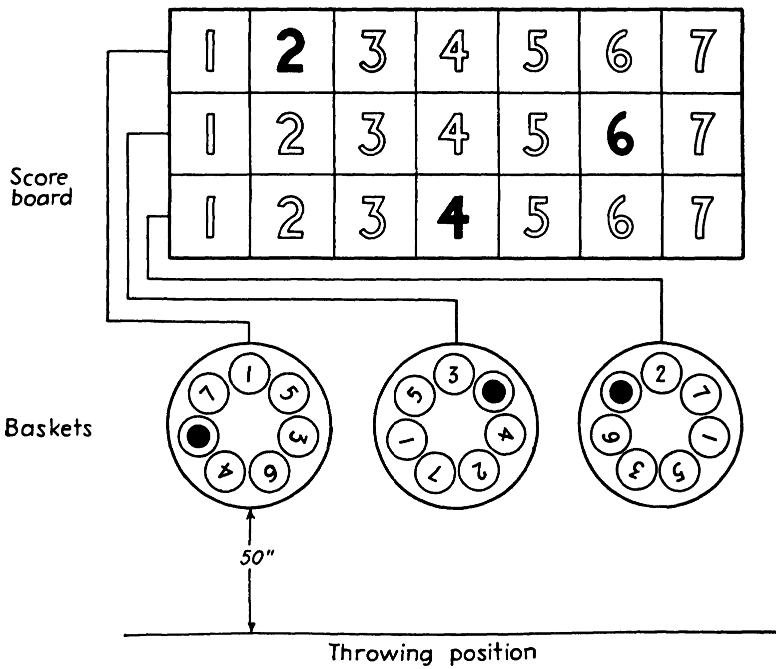

Fig. 21.—The apparatus.

wire-mesh baskets, flared toward the top, each equipped with a wooden bottom containing seven pockets. The pockets, numbered from 1 to 7, each contained a switch controlling a corresponding number on a board above the baskets.

The baskets were arranged in a row under the score board and separated from each other by intervals of three feet. The players sat or stood behind a long table, in a playing position 50 inches from the baskets. Basket tops were 40 inches from the floor, the bottoms 20 inches lower.

The players threw solid rubber balls, $1\frac{1}{4}$ inches in diameter,

into the baskets in an effort to hit certain numbers. The numbers tried for were designated before the throw by placing metal disks of various denominations on numbers printed on a green oilcloth at each playing position.

The Play

A ball thrown, at random or otherwise, into one of the baskets is certain to operate a switch and light a number in the row on the scoreboard belonging to that basket. The opportunity for the exercise of skill exists in the fact that a ball properly thrown will operate the number previously designated. By the terms of the play printed on the green oilcloth in front of him, a player may elect to play only for odd or even numbers (the number 7 losing to the management of the game), receiving an amount equal to his wager when he wins. He may, if he chooses, try for any number from 1 to 7, inclusive, winning at 5 for 1. He may elect, instead, to play one ball into each of the three baskets, or join with one or two other persons in playing into the baskets, trying for a combined score of 3 to 21, inclusive, in the three baskets, at increased odds depending upon the theoretical frequency of the combination tried for.

Theoretically, by pure chance, the sums 3 and 21—three 1's and three 7's—appear once each per 343 random throws. By the legend on the green table they win at 100 for 1. Single numbers each appear once in seven throws, on the average, provided there is no bias due to position in the baskets.

Consistent winning at the game depends upon the development of such skill that the odds offered by the management may be overcome. The odds offered, however, do not parallel the theoretical frequency of the numbers that may be tried for, some combinations offering better contracts to the players than others. Table XIII, page 105, shows the distribution of these sums according to the laws of chance and, paralleling them, in the last two columns, the distribution of the odds offered by the management of the game. From a study of these distributions, the problem confronting the players should be clear. The sums at the ends of the distribution offer poor contracts. The sums from 9 to 15, inclusive, offer the best contracts, since the paying odds are nearer the percentage frequency of appearance of these

combinations. Thus, on the 12, which appears once in nine trials, on the average, the management pays 6 for 1. To break even, the player would have to make this combination appear once every seven trials. If a player could make this number appear once every six or fewer trials, he could win against the

TABLE XIII.—THEORETICAL VS. PLAYING ODDS

Sum of numbers, three baskets	Frequencies of sums	Percentage of totals	Playing odds, odds to 1	Percentage odds
3	1	.29	100	1.00
4	3	.87	50	2.00
5	6	1.75	25	4.00
6	10	2.92	15	6.67
7	15	4.37	12	8.33
8	21	6.12	10	10.00
9	28	8.16	9	11.11
10	33	9.62	8	12.50
11	36	10.49	7	14.29
12	37	10.79	6	16.67
13	36	10.49	7	14.29
14	33	9.62	8	12.50
15	28	8.16	9	11.11
16	21	6.12	10	10.00
17	15	4.37	12	8.33
18	10	2.92	15	6.67
19	6	1.75	25	4.00
20	3	.87	50	2.00
21	1	.29	100	1.00
	343	99.97		

INTERPRETATION: The sum 12, in the three baskets, occurs 37 times by chance along among 343 trials It thus occurs 10 79 per cent of the trials and wins at the rate of 6 for 1, but, to break even, it would have to be made to appear 16.67 per cent of the trials. (This assumes continuous betting on the number 12.)

odds. By chance alone, *i.e.*, random throwing, without skill, he would lose in the long run, of course.

On the basis of chance alone, the management stands to win quite handsome profits from the game. If the players could develop sufficient skill to overcome the odds consistently, the management could also lose quite heavily, since the players are competing against the house odds and not against each other. If two or more players are permitted, without collusion among

them, to throw into the same basket in an effort to make different numbers or combinations, they may nullify each other's skill and reduce the scores to their theoretical expectancy.

Hypotheses

Inspection of the apparatus and observation of the methods of play led to the conclusion that attainment of skill by the players was reasonably certain; that the game was adapted to individual, cooperative, and competitive play; and that it had merit as a device for affording recreation and amusement.

The following hypotheses, later tested by the experiment, were formulated:

1. That random throwing should yield approximately equal percentages of 1's, 2's, 3's, etc., according to the laws of chance.
2. That effort and practice by the player should produce a substantially higher proportion of a certain number than would be obtained by chance alone.
3. That learning curves plotted from the data of practice should correspond generally to the known curves of muscular learning.

To test these hypotheses, it was proposed that an experiment be conducted with unskilled players under controlled conditions for a period of several weeks, carefully observing and tabulating the results, and comparing the latter with known characteristics of skilled performances.

The Procedure

Three persons in their early thirties, a woman and two men, were selected as subjects. None of them had seen the game before this time, but all had played games of chance and skill at amusement parks. Throughout the experiment these three manifested a keen interest in trying to improve their skill.

The trials were carried out on three evenings each week for three weeks. Practice began at 8:00 P.M. The practice for each evening, during the first six evenings, consisted of five periods of 15 minutes each, separated by 5-minute rest periods. The seventh evening was devoted to competitive play for a predetermined combined score in the three baskets. The eighth

evening was devoted to four practice periods for individual scores and one period of random throwing. The ninth evening was devoted to three periods of 100 trials each, in which the players tried cooperatively for a predetermined total score. Only the results of the individual play are reported here.

Analysis of Results

1. *Random Throwing.*—Study of the results of random throwing revealed no bias toward any one or several numbers in the baskets. The distributions followed the theoretical expectancies so closely that only two periods were devoted to this test.

2. *Individual Throwing for a Predetermined Score.*—Thirty-two periods of 15 minutes each were devoted to this part of the experiment, each player throwing one ball into the basket assigned to him, and all throwing at approximately the same time. After each throw was recorded, the apparatus was reset for the next throw and a signal of readiness given the players. Each player threw 4,008 times during this part of the experiment.

Throughout the experiment the players were permitted to choose their own style of throwing, but the numbers to be tried for were designated for each 15-minute period by the experimenter. This designation of numbers was necessary in order to cover all the holes in the baskets, since some were much more difficult to hit than others.

Between the third and fourth evenings, the positions of the baskets were changed by rotating them through an arc of 180 degrees. This second position was maintained for the remainder of the experiment.

The results of this series of trials are shown in terms of percentage variation from theoretical expectancy in Figs. 22, 23, and 24, pages 108, 109, and 110, respectively, and the combined scores for the three players in Fig. 25, page 110. The upper graph in each figure shows the detailed results for each 15-minute period in sequence. The lower graph shows these data smoothed by the method of the five-point moving average. The zero line is the level of theoretical expectancy, the level that should be attained by pure chance. The departures above this line indicate the effect of skill.

By chance alone each player should have scored the numbers tried for one-seventh of the total number of throws, or 572 times. All exceeded this figure: Mrs. L with 842 by 47 per cent; Mr. C with 896 by 57 per cent; and Mr. L with 622 by 9 per cent. These are the average percentages by which the curves depart from the theoretical expectancies. Clearly Mr. C was

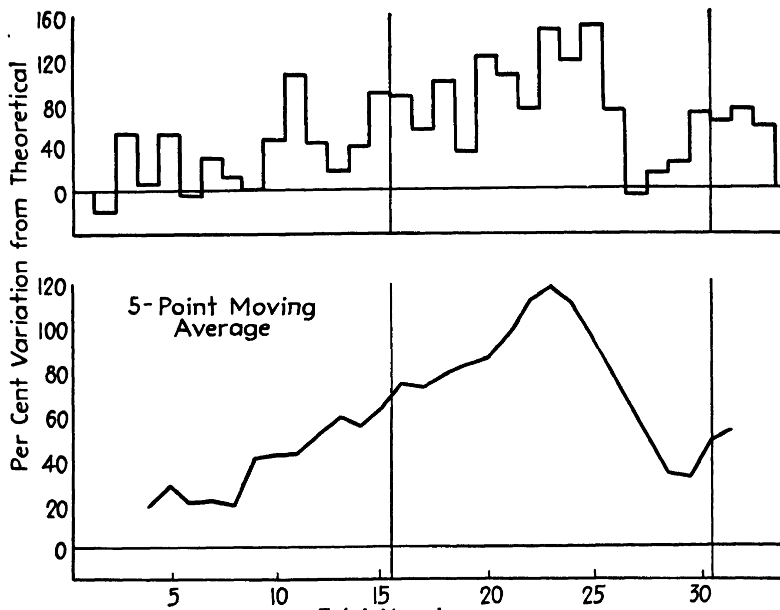

FIG. 22.—Learning curve: Mr. C.

the most proficient performer, with Mrs. L a close second, and Mr. L a poor third.

The average performance of the three players together is 31.5 per cent better than chance. This is a level of performance sufficient to offset the management's odds and to win moderately for the players. For Mr. C, with his average performance 57 per cent better than chance, the returns would be 134 for every 100 units wagered. Mr. L would lose in the long run, however, since his average performance of only 9 per cent better than chance is not good enough to offset the management's odds.

The following characteristics may be noted in these performances:

1. Very few of the scores fall at, or below, the theoretical expectancy. (Some of these were undoubtedly due to the fact that the numbers thrown for were in a part of the basket difficult to hit, something the player could have avoided had he been free to choose his own number.)

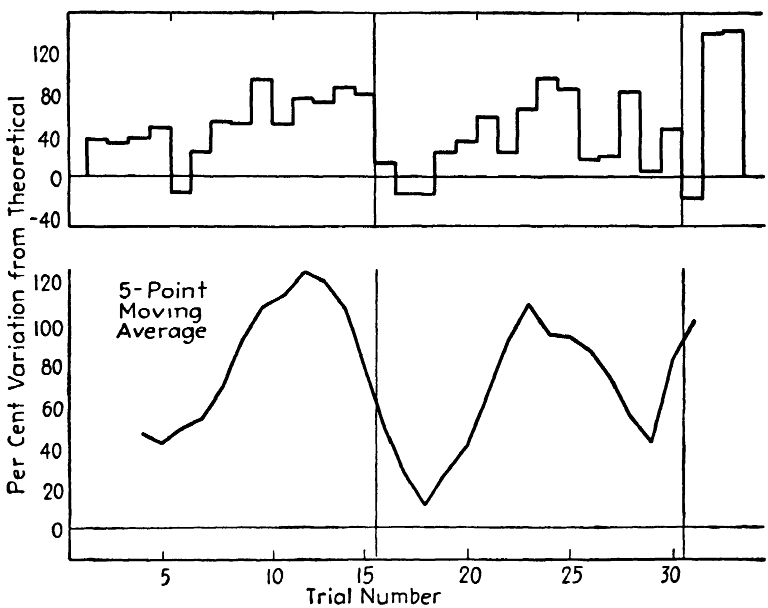

FIG. 23.—Learning curve: Mrs. L.

2. In general, the players improved their skill during each evening's play.

3. The play on later evenings of the week showed improvement over the play in the earlier part of the week.

4. Loss of skill occurred over the weekends, but such skill was soon recovered when play was resumed.

5. Marked differences of ability among the players are revealed in the relative heights and trends of the graphs. This corresponds with the known facts of individual differences in skills.

These results tend strongly to confirm the second and third hypotheses. The fact of marked individual differences in the

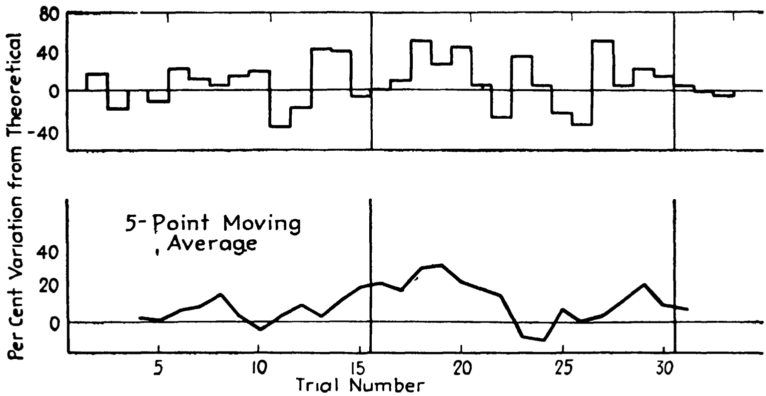

FIG. 24.—Learning curve: Mr. L.

results might have been anticipated in the hypotheses, since their presence as revealed in the data tends to support the gen-

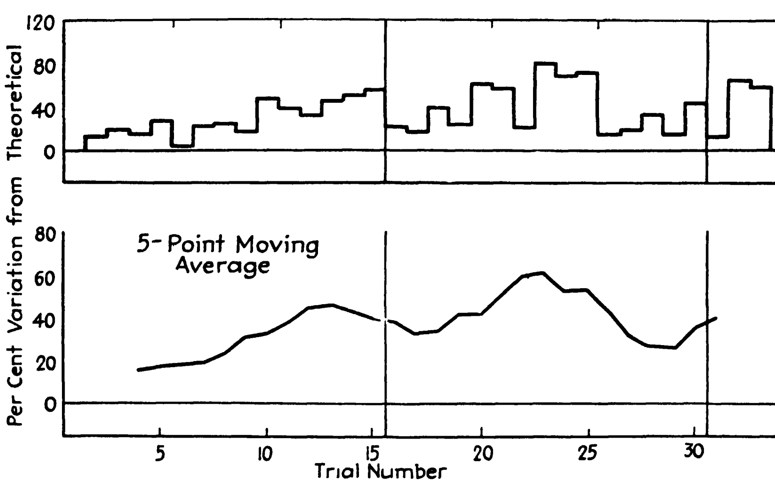

FIG. 25.—Learning curve: combined scores.

eral assumption that scores in the game will be influenced by individual skills.

The curves in the foregoing figures reflect the two main criteria of skilled performances. The first of these is that there shall be a significant difference between the distribution of results obtained when effort is directed toward the prescribed end, thus creating the situation favoring the development of skill, and the distribution that results from random, undirected effort. The other is that there shall be derived from successive trials one or more curves that follow, in general, the typical learning curve, *i.e.*, rapid initial rise, gradual negative acceleration or flattening out, and approach to or attainment of a maximum beyond which no further improvement is indicated—the physiological limit.

TABLE XIV.—COMPARISON OF EARNINGS

Number	Class of employees	Wages, per cent of average	Average length of service, months
13	Old employees, passed tests	116.1	32
10	Old employees, failed tests	93.6	24
16	New employees, passed tests	97.4	2
9	Trainees, passed tests	113.3	½

Another characteristic, most important for industry, is also shown in this study. This is the wide differences in rapidity of learning the skilled performance and the similarly great differences in final level attained among the three persons. This is directly analogous with the typical situation in industry. The desirability of initial selection of those who can learn both rapidly and to a high final level is evident. A study of the several learning curves should be enlightening to anyone who is concerned with the nature of skilled performances.

While we have no similar curves for industrial jobs showing the rate at which learning proceeds for those who have been selected by test as against those selected by interview or rejected by test, Table XIV, above, gives the achievements in terms of piece-rate earnings of four different categories of employees against the average wages paid in the department. Both groups of old employees had been selected by interview, whereas the new employees and the trainees had been selected by test

as well as interview. It may be noted that the new employees were earning higher piece-rate wages, on the average, after 2 months of service than the older employees who failed the test and who averaged 24 months of service. It is also notable that the trainees after 2 weeks of service were earning wages almost equal to the average of the old employees who passed the test and whose average length of service was 32 months. Selection by test and specific job training are clearly justified by such results.

CHAPTER XII

MEASUREMENT OF ACCIDENT-PRONENESS

Until quite recently it was customary among many engineers and psychologists, as well as among the lay public, to look upon industrial and traffic accidents as primarily due to chance factors which could not be brought under control. More recently there has been a marked tendency to ascribe many accidents to personal behavior in the individuals concerned, *i.e.*, to variations from a standard pattern of right and safe behavior. The remedy obviously lay in training, since the deviations were assumed to be matters of habit that were largely under voluntary control.

The effectiveness of extensive and intensive safety training programs instituted by various industrial organizations has often been startling. In some other instances such programs seem to have almost reached the limits of effectiveness, considering the time and effort devoted to them, but accidents continue, although at a lower rate. Fortunately, however, the means of control have not been exhausted.

EARLY FINDINGS

Investigations by Greenwood and Wood [1] in 1919, followed by that of Newbold [2] in 1926 covering 16,000 British cases, and further supported by the findings of Bingham and others in the United States,[3] point to one conclusion: *Some individuals are, by reason of certain inherent characteristics, predisposed*

[1] M. GREENWOOD and H. M. WOOD, The Incidence of Industrial Accidents, *Industrial Fatigue Research Board Report,* no. 4, 1919.

[2] E. M. NEWBOLD, A Contribution to the Study of the Human Factor in the Causation of Accidents, *Industrial Fatigue Research Board Report,* no. 34, 1926.

[3] W. V. BINGHAM, Individual Differences in Industrial Personnel, *Eugenical News,* vol. 15, no. 2, February, 1930.

to have accidents. It was upon such a conclusion that the experimental investigations of Farmer and Chambers [1] were predicated.

These latter investigations, made under the auspices of the British Industrial Health Research Board (formerly the British Industrial Fatigue Research Board) and published in its series of bulletins, demonstrated conclusively the possibility of further reduction of accidents through the identification of individuals having the accident-proneness characteristics. Their results indicated that the selection of industrial workers by tests would effect a 50 per cent reduction in accidents, since it has been amply demonstrated that the most able workers are also the safest and that work done correctly is done safely.

The British investigators found no relationship between intelligence, as usually measured, and the accident records. They did find, however, that three tests of the considerable battery used showed substantial relationships with the records. These three tests were Choice Reaction Time, Dotting, and the Pursuit Meter, all well known to psychologists.

It is important to note that each of the three tests involves in its performance the use of attention and perception in conjunction with muscular movements. Since it was quite impossible to separate the perceptual factor from the motor control factor, the two entered into the final scores in unknown amounts. The effect of this was to reduce the amount of actual relationship that was present in the data.

Relationships between Tests and Accidents

A more limited investigation [2] made some years ago by this writer indicated that only moderate relationships between accident records and tests were obtainable when scores were handled in the conventional method used by the British investigators.

[1] E. Farmer, and E. G. Chambers, A Study of Personal Qualities in Accident Proneness and Proficiency, *Industrial Health Research Board Report,* no. 55, p. 59, 1929.

[2] C. A. Drake, The Prediction and Control of Accidents, *The Scientific Monthly,* July, 1940, pp. 74-76; Accident-proneness: A Hypothesis, *Character and Personality,* June, 1940, pp. 335-341. See also, Detecting the Accident-prone Worker, *Personnel,* March, 1942, pp. 276-281.

Far more significant relationships were found when the differences between scores on perceptual and motor tests were compared with an index that took account of both frequency and severity of accidents.

Four of the tests used in the original battery were planned as the result of a considerable number of job analyses and work observations made in one department of a factory and reported in detail elsewhere.[1] They consisted of two motor manipulation tests and two inspection or visual perception tests. A fifth test, the O'Connor Tweezer Dexterity Test, was added for purposes of comparison, as the other tests were specially designed and still unstandardized.

All the tests used were of the work-limit type, *i.e.*, they consisted of a definite number of pieces to be manipulated, and the scores were determined by the time taken to accomplish the tasks set. To facilitate direct comparisons among the several tests, all raw scores were converted into standard scores, with a mean of 50 and a standard deviation of 10 on the bases of the distributions of scores made by the several hundred persons tested.

The motor tests involved principally dexterity of the fingers, together with some motions of the wrists and arms. Pin Board E (Fig. 8, page 61) required the testee to pick up round pins, ⅛ inch in diameter and 2 inches in length, one with each hand, and to place them simultaneously in parallel rows of holes in a large wooden tray. The Right-right Turning Test involved the turning of 10 pairs of machine screws into threaded holes in a vertical steel plate. The O'Connor test required the testee to work with only one hand in placing small metal pins, with the aid of tweezers, in parallel rows of vertical holes in a wooden block.

The inspection tests were primarily tests of visual perception, but both involved some motor manipulation in handling the parts. The Spiral Inspection Test (Fig. 4, page 41) consisted of 100 small aluminum spirals, each of which had been punched with a small hole near one end. Fifty were punched "standard" —2½ turns from the end. The remaining 50 were punched

[1] C. A. Drake (with H. D. Oleen), The Technique of Testing, *Factory Management and Maintenance*, vol. 96, no. 3, pp. 71-78, March, 1938; C. A. Drake, Inspection for Inspectors, *American Machinist*, vol. 82, no. 17, pp. 766-768, Aug. 24, 1938.

"off-standard," *i.e.*, at various distances other than the standard. The testee was required to separate the standard items from the others. The Case Inspection Test (Fig. 5, page 41) introduced an additional perceptual operation. One hundred and twenty colored metal pencil cases were to be sorted into six compartments of a tray by colors, but 30 defaced (punch-marked) items were to be detected and placed in a separate compartment.

The experimental group consisted of 40 female operators from one metalworking department of the factory selected by the foreman in response to this request: "Some of the best, some of the average, and some of the poorest." The range of scores on each of the several tests indicated that this had been conscientiously done. Cards bearing the names of these 40 operators were subsequently handed to this foreman on two different occasions about three weeks apart with the request that he rank these operators in order of efficiency. Each time he arranged the cards without hesitation, his two rankings agreeing with each other .9. We later learned that these rankings had been conditioned by, and agreed very well with, the operators' piece-rate earnings.

These 40 operators had 73 accidents during the 17 months covered by the accident records, against 71 accidents recorded for the other 39 operators in this department. Twenty-three of the selected group of 40 operators suffered from one to nine accidents each during the period, against 23 in the remaining group of 39 who had from one to fourteen accidents each. From this and other data we concluded that the experimental group was a representative sample.

Since the accidents recorded differed in severity, and since 20 out of 40 operators had been employed less than the 17 months covered by the records, an "accident index" figure was computed for each person by the formula:

$$\text{Accident index} = \frac{\text{number of accidents} \times \text{severity}}{\text{length of service in months}}$$

Average severity was estimated by the registered nurse in charge, using a scale from 1, slight, to 10, very severe. Severity of each accident was largely dependent upon the number of redressings required rather than upon the extent of impairment in working

ability. For the 20 operators whose period of service exceeded the 17 months covered by the records, the figure 17 was used in the denominator. A high index figure, therefore, represents many accidents, or severe accidents, or both.

Systematic statistical and graphical examinations were made of the direct relationships between the accident index and each of the several sets of test scores without arriving at any significant conclusions. Psychographs, of the type shown in Figs. 26*a* and 26*b*, below, were also made for the operators, and the

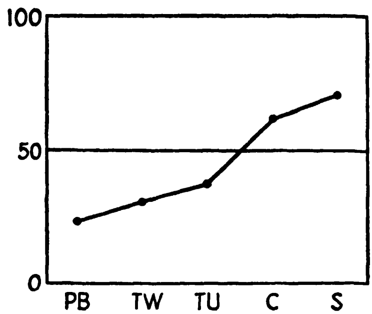

FIG. 26*a*.—The "safe" type; accident index = .00.

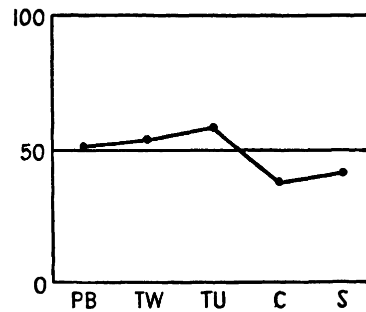

FIG. 26*b*.—The "accident-prone" type; accident index = 1.00.

accident index figures were inscribed in the upper right-hand corners.

During a somewhat casual manipulation of these psychograph cards, it was observed that the accident index figures tended to be high when the scores on the motor tests were higher than the scores on the perception tests. It was also observed that the accident index figures were low, or zero, when the scores on the perception tests were higher than the scores on the motor tests (see Fig. 26*a* and 26*b*).

In following up this promising clue, the numerical scores were again examined in relation to the accident index figures, this time by subtracting each test score from every other test score for that individual and computing the relationships between such differences and the accident index. Computations were also made of the relationships of the differences among various groups of scores and the accident index.

From the foregoing computations it was concluded that the set of differences resulting from the subtraction of the scores on

the Right-right Turning Test from the scores on the Spiral Inspection Test exhibited the most significant relationships with the accident index figures. These relationships are shown graphically in Fig. 27, below. The probability that this differentiation is valid and not the result of chance is high.

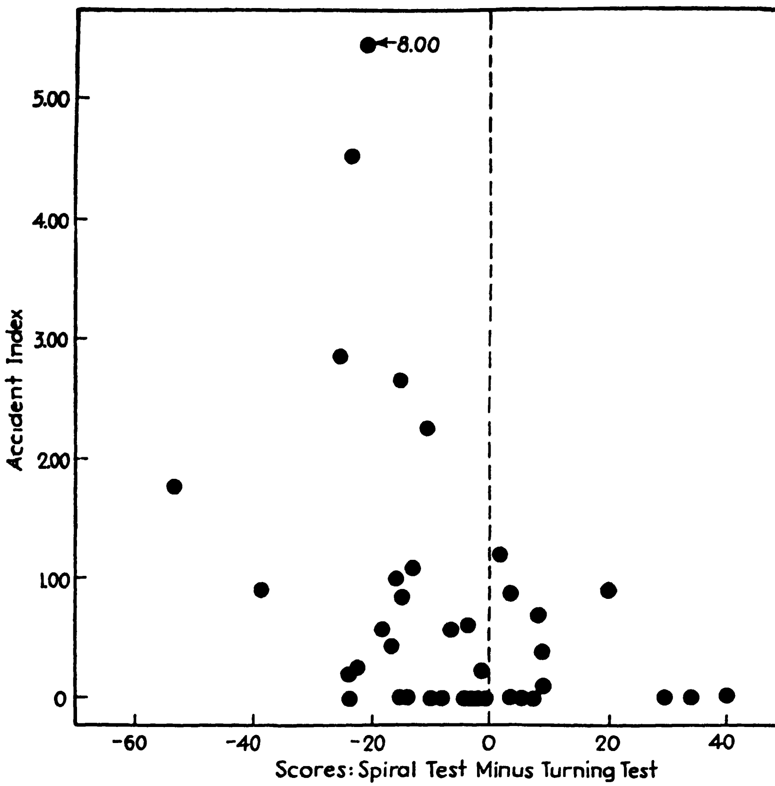

FIG. 27.—Comparison of accident index with test scores.

The average difference-score for the 17 accident-free persons in the group is -2.9, with a standard deviation of 14.1; for the 23 accident cases the average difference-score is -12.9, with a standard deviation of 14.6. The difference between these means is 10.0, and, by the usual formula, the standard deviation of the difference is 4.5. The difference is therefore 2.2 times its standard deviation, indicating that the probability is 98.6 out of 100 that the difference is genuine (greater than zero).

A comparison of Fig. 27, page 118, with similar diagrams on which the scores of each of these two tests are plotted against the accident index clearly reveals a lack of significant linear relationship. This probably accounts for the failure of the earlier comparisons to disclose the existing relationships.

That these two tests are measuring quite different abilities may be inferred from a study of the manipulations called for in the performance of the tests and from their low intercorrelation. The latter is of the order of .2, as computed from the whole group tested later—operators and applicants. Each gives an approximately normal distribution of scores, ranging from 3.61 to 11.68 minutes for the Spiral Inspection Test, and from 4.42 to 12.31 minutes for the Right-right Turning Test, in terms of raw scores. Retest reliabilities are .8 for the former and .9 for the latter, both uncorrected.

It may be inferred that the Spiral Inspection Test is measuring visual perception, primarily, and that the Right-right Turning Test is measuring a manual or motor speed and control factor. These are not the best tests that can be devised for measuring such factors. The Spiral Inspection Test involves too much perception in only two dimensions and too much manipulation, this latter overlapping what is measured by the Right-right Turning Test. This latter test involves too much positioning, or fitting of the screws into the plate (to get them started), a factor quite different from the mere turning of the screws, thus not affording a pure measure of the apparently more significant turning ability.

Further research in this field would seem to require several tests of perception in two and three dimensions, uncomplicated by manual and other disturbing factors, and several tests of manual or muscular manipulation, as nearly free from perceptual and other complicating factors as it is possible to devise them.

A Hypothesis

The foregoing results lead to the hypothesis that accident-proneness is a phenomenon associated with discrepancies in level between perception and motor reaction.' It was observed that persons whose perceptual level is equal to or higher than their motor level are relatively safe, while those whose perceptual

level is lower than their motor level are accident-prone, with records of more frequent and more severe accidents than the former group. This implies that those who can see faster than they can react are relatively safe, while those who react faster than they can see are accident-prone.

It does not follow, however, that all uncorrected defects of vision contribute directly to slow perceptions and thus to accident-proneness. It is apparently possible to have a variety of such defects and still get perceptual cues that are quite adequate for effective and safe behavior. Although defective vision tends to reduce the speed of perception, the level of speed of perception is not necessarily reduced below the level of speed of the motor reactions.

There is some evidence to indicate that certain individuals suffer an early and rapid breakdown of the speed of perception under the influence of fatigue, alcohol, and distraction. In others the breakdown seems first to affect the speed of motor reactions, leaving the speed of perception relatively unimpaired. This latter group would tend to be relatively safe under these influences as long as the perceptual level is not reduced. On this hypothesis, however, the introspective evidence of the individual himself should not be sufficient to justify his working at a dangerous task while excessively fatigued, or driving his automobile while intoxicated. The foregoing evidence does raise some doubt as to the validity of measures of vision and of the alcoholic content of the blood as bases for a judgment of impaired skill and safety. The effect of these factors on perception seems to be the real criterion.

It is significant that in the writer's investigation two relatively safe groups of industrial workers were found—the very good piece-rate workers on assembly operations and the very slow group of day-rate workers. It would seem that the latter avoided accidents by a slow rate of work. This is shown in Table XV, page 121. The preponderance of accidents in the third quarter of the group—the quarter containing the workers who were barely making the piece rates—is striking. If tests had been used in employment selection, most of the persons in these two lower quarters would not have been employed. This alone would have effected a reduction of accidents amounting to 55 per cent, a result that compares favorably with the estimate

made by Farmer and Chambers. If the difference-scores had also been used for selection, the reduction would have been much greater.

TABLE XV.—RELATION OF ACCIDENT INDEX TO EARNINGS

Foreman's ranks (earnings)	Sum of index figures, by eighths	Sum of index figures, by quarters
1–5	2.24	4.83
6–10	2.59	
11–15	2.01	7.28
16–20	5.27	
21–25	4.49	13.35
26–30	8.86	
31–35	1.88	7.27
36–40	5.39	

TABLE XVI.—ACCIDENT INDEX AND TEST DIFFERENCE-SCORES

Difference-scores (perception minus motor)	Sums of accident index by		
	Eighths	Quarters	Halves
+60 to +29	.71	4.29	7.71
+23 to +18	3.53		
+15 to + 8	.57	3.42	
+ 6 to − 5	2.85		
− 6 to −11	2.88	6.57	25.04
−12 to −15	3.69		
−15 to −21	9.00	18.47	
−21 to −48	9.47		

Table XVI, above, shows, for the same group, the marked tendency of the difference-scores to pick out the individuals with the high frequencies and severities.

If employee selection were made using both methods of test score application, a reduction in the average accident index of as much as 75 per cent might be achieved, and this quite independently of other safety education and training measures. In fact, in one group of new employees thus selected, the actual reduction was 70 per cent, in comparison with the records of a group selected by the usual interview technique.

A still further reduction in accidents might be achieved by classifying jobs and work areas in terms of the hazard they present and assigning to the most dangerous work the applicants least likely to have accidents, and to the least dangerous work the applicants most likely to have accidents. By this method a given set of workers could be distributed so as to effect a substantial reduction in accident rate and severity.

Implications

What are the prospects for attacking this apparently inherent accident-proneness in some persons by education and training? It must be confessed that the results of the attempts to modify the test scores by training have been discouraging. The tests seem to be measuring innate qualities that can be modified only slightly, and that with great difficulty and perhaps only temporarily.

The better approach to control seems to be to discover the degree of accident-proneness of a person and then to place him in an environment in such a way that he will not be subjected to hazards beyond his tested limitations. Knowing his limitations, he may, by the exercise of judgment, avoid situations that might have disastrous results. If the hypothesis is correct, and if the measurements represent a new area of stable individual differences, the airplane pilot, the automobile driver, and the dish-breaking domestic servant are as suitable subjects for testing as is the industrial worker. One of the important characteristics of the proposal is that, since the measurements can be made before selection and before training is undertaken, many severe and some fatal accidents may be avoided.

CHAPTER XIII

HOW TO CONSTRUCT TESTS

Among the readers of this book there will certainly be some who will wish to design tests for their own experimentation. Only by actually designing and applying such tests can one gain real insight into the problems involved or develop an appreciation of the advantages of the measurements secured.

A Simple Inspection Test

For purposes of illustration, let us take a simple industrial job on which the operators are required to inspect the product, one unit at a time, at a standard rate. Let us say that this job involves the examination of a round lead pencil to determine whether the lead is correctly centered in the wood cover. In this operation, there is obviously a manipulative factor involving pickup, transport, positioning at the point of observation, transport again, and release. This operation cycle requires a separate test of motor abilities involving finger movement, wrist movement, forearm movement, and, to some extent, full-arm movement. The design of such a test can be left for later consideration. The immediate concern is with a test of the ability to perceive small differences between the circumference of the cover and the circumference of the lead; or to perceive differences in distances between the center of the cover and the center of the lead; or, if one prefers another expression, to perceive a difference between symmetry and nonsymmetry.

On the job the operator is required to reject pencils that differ by more than some standard amount from perfect concentricity. That is, the operation involves the ability to detect small differences by means of visual perception.

Mental Operations

In considering the mental operation involved, it seems relatively easy to design a suitable pencil-and-paper test. The test items will start with figures as far from concentricity as the worst pencil to be inspected, and will approach, in successive figures, the limit of perfect concentricity. The results of the first attempt to design such a test are shown in Fig. 28, be-

Fig. 28. —Original Visual Perception Test.

low. This test was designed by the writer's associate, Holger D. Oleen, to meet a situation similar to that described. The first items on this test are so easy that anyone should be able to distinguish the ones to be rejected. The last units are so near perfect concentricity that very few persons can mark them with certainty. In fact, at a point somewhere before the last line is reached, the average person has started to make a considerable number of errors. Some of these errors are due to careless observation, but, as the difficulty of the test increases, the errors or omissions tend to increase rapidly. At some point the test goes beyond the place, called by psychologists the *limen,* at which the individual is able to discriminate the nonconcentric items from the concentric ones.

The test provides for a time score, as well as an error score, since individuals are known to differ widely in the speed of their perception. In fact, in practical inspection work in the factory, the speed with which the inspection is performed is quite as important as the ability to discriminate the quality of the units.

Reliability and Validity

Every constructor of a test must eventually answer in the affirmative two questions to justify the use of his test in the serious business of employee selection. The first is, "Is the test reliable?" *i.e.*, does it give identical or very similar results every time it is applied to the same persons? The other is, "Is the test valid?" meaning, does it really measure the ability called for on the job for which it is designed?

In answering the first question it must be remembered that reliability in a test is largely dependent upon the length of the measure, *i.e.*, upon the amount of time devoted to its performance. Merely by increasing the number of cycles to be completed we can increase this reliability. The best length for practical application will always be a compromise between the ideal of perfect reliability attainable at an infinite amount of time and the unsatisfactory reliability that results when too little time is devoted to the test performance.

In evaluating the technique explained in this book, we object to the usual appraisal of both reliability and validity in terms of a coefficient of correlation. We have emphasized that we are concerned primarily with the elimination from further consideration of those persons least suitable for the jobs in question, and with the identification of the best for purposes of further selection by interview, physical examination, etc. A simple percentage measure of the proportion of the same individuals who fall in the "best" groups on the test scores derived from successive applications of the test is therefore considered a better measure of reliability. Similarly, the percentage of those in the selected group who make good on the job is a better measure of validity.

Critical Score

In the preliminary tryout of the test on experienced operators, it is desirable to find the point, or score, above which falls the part of the distribution containing the old employees who are predominantly successful on the job and below which is a large proportion of unsatisfactory or incompetent employees. If the test used is really measuring basic abilities, new employees chosen from among those who had scores above this point should, in the vast majority of cases, become workers quite as satisfactory on the average as the old employees who scored above this level. The group scoring below this point should, and usually will, show about the same proportion of unsatisfactory and incompetent employees as was found among the group of older workers.

In the practical work of selection, the choice of a suitable critical score will depend in part on the condition of the labor market. When applicants are plentiful, selection can be rigorously applied. When applicants are scarce, however, this critical score, or point of acceptability, will tend to be lower as a matter of expediency.

It must be remembered that if we skim off the best 5 per cent of applicants, practically all of them will succeed conspicuously on the job. If we must skim off the best 10 per cent, owing to the condition of the labor market, the selection procedure will still give us highly satisfactory employees. If, however, we find it necessary to take the best 50 per cent, or more, of applicants, many of them will be near the average, and the selection procedure will not show that percentage of success that will be obtained with the more rigorous selection of the best 5 or 10 per cent.

The time has now passed when it is economically or scientifically justifiable to give such tests experimentally and to hire by interview without regard to the test results, the assumption being that these employees will then be followed up in their subsequent employment to determine the value of the test technique. On the basis of experience, it is recommended that where lack of a better method of setting a critical score exists, the average score be taken as the critical score and that no

person be employed for the jobs to which the test is related whose score is below this average. This will usually give positive assurance that practically all of the slow and incompetent will be eliminated at the outset.

A Simple Dual Test

Let us now consider an industrial job on which the worker picks up two small parts simultaneously from hoppers to the right and left of the work space, transports them to a point in the middle plane of the body, and drops them simultaneously into a small carton. After a number of these cycles, the carton is closed by the action of both hands and pushed to the side of the work area with one hand, while the other hand sets a new carton in position in the holding fixture in front of the operator. This is a simple repetitive operation of a type very common in industry. It belongs to the category of dual operations, since both hands perform the task simultaneously with two identical parts.

By direct observation, without the aid of time and motion analyses, it is obvious that this cycle of operations consists of the following five elements: (1) pickup; (2) transport; (3) position; (4) release; and (5) return empty. The most variable element among these five will be that of positioning. Also, it is the most important. The design of the test should embody the five elements approximately in the same proportions they occupy in the work cycle. If the positioning element occupies 20 per cent of the cycle on the job, it should be given at least this proportion of the cycle in the design of the test.

The presentation of the parts or test items in the hoppers on the test should parallel their arrangement in the hoppers on the job. That is, if the parts are pre-positioned in the hoppers in the latter instance, the test should be arranged accordingly. In general, speed of operation can be increased by pre-positioning.

The same analogy holds true for positioning. The positioning element on the test should be similar to, and at least as difficult as, the positioning on the job. If indexing and positioning are both present in the cycle, they must be presented similarly in the test. It must also be remembered that positioning is largely muscular and kinesthetic, with only a minor dependence upon

FIG. 29.—Textile Thread Test.

visual perception, while indexing is usually heavily dependent upon visual perceptual ability. The job may often require a separate test of this latter ability. Evidence to support this observation may be deduced from Table II, page 27.

A simple dual test of the kind shown in Fig. 19, page 82, should be adequate for this job. A simple pin board, using pins of a size as small as the items used on the job, should also answer the purpose.

With a test of this short-cycle variety, it will be found that testees will, in general, very soon attain their maximum speed. This characteristic implies quick learning on the job and indicates that a forecast of production based upon the test scores will be more accurate than is the case with longer cycles.

A somewhat different type of test, involving many repetitions of a very simple cycle, is shown in Fig. 29, page 128. The testee is required to manipulate a string, using both hands to insert it in a series of "pigtail" guides. This operation parallels certain jobs in the textile industry. The manipulation involves finger, wrist, and full-arm dexterity, while the spacing of the test units imposes a handicap upon persons with large, clumsy hands. The vertical mounting of the test places it in the usual plane of operation used on the job, frequently an important consideration.

Special Considerations in Design

In the design of dexterity tests it is most important to require that the muscular movements involved in manipulation be as small or as refined as those required on the job, and the items to be manipulated should be at least as small as the parts actually handled in the work. Selecting assemblers who position and operate an electrical screw driver using a full-hand grasp and arm and wrist motions by means of a small pin board test is manifestly absurd.

Even more absurd, but unfortunately not unknown, are attempts to select assemblers of electrical instruments who follow a prescribed plan quite outside their own initiative by means of so-called tests of mechanical insight or ingenuity. A simple analysis of the finger and other motions called for on such a job would easily give the requirements for an appropriate test.

It cannot be too emphatically asserted that one dexterity

found in satisfactory amount in an applicant is no guarantee of the presence of other dexterities. There is no guarantee that high general hand dexterity, for example, assures the presence of even mediocre hand-foot coordination or of dual dexterity. The need for designing a separate test for each of the types of ability required by jobs and groups of similar jobs should be apparent from a study of Fig. 30, page 131, showing the test profiles of 10 industrial employees.

Interpretation of Results

In Fig. 30, No. 1 represents a worker much above average on all of the abilities tested. Number 2 is at the other extreme, uniformly poor on all the tests. Where a management is not "sold" on testing, or where lack of systematic handling and insight may assign workers without careful regard to these indicated abilities and handicaps, the only safe procedure is to select persons who are, like No. 1, high on all abilities—but not many of these are to be found. Numbers 3, 4, 5, and 6 all have about the same level of general hand dexterity, but note that No. 4 has a much higher level of inspection ability and intelligence than No. 3, whereas No. 5 has a very high level of hand-foot coordination ability, and that this latter ability is most notably lacking in No. 6. Numbers 7 and 8 present almost diametrically opposite patterns of abilities, except for general hand dexterity, in which they are equal. Numbers 9 and 10 present somewhat similar profiles of ability, with this exception. In hand-foot coordination No. 9 is deficient and No. 10 excels, whereas in inspection ability the situation is reversed, *i.e.*, No. 9 is superior and No. 10 is markedly deficient.

Profiles of this sort are extremely useful in assigning men to jobs when the tests are appropriate for such jobs. In some instances they might be quite misleading, for example, if a hand-and-foot job required use of both feet, and the test was a two-hand and one-foot test. While the latter test would probably give a better basis for assignment than any of the other tests, a man who scored well on it might have a special disability affecting the use of the foot that was not involved in the test but that would have to be used on the job.

The practical use of such profiles is to enable the person making assignments to select the applicant possessing the combination of abilities called for on the job. Thus, a job requiring superior ability in both hand-foot coordination and in inspection ability would obviously be filled by No. 5. A job requiring general hand dexterity and superior intelligence could be filled by several. But more than half of these employees have special dis-

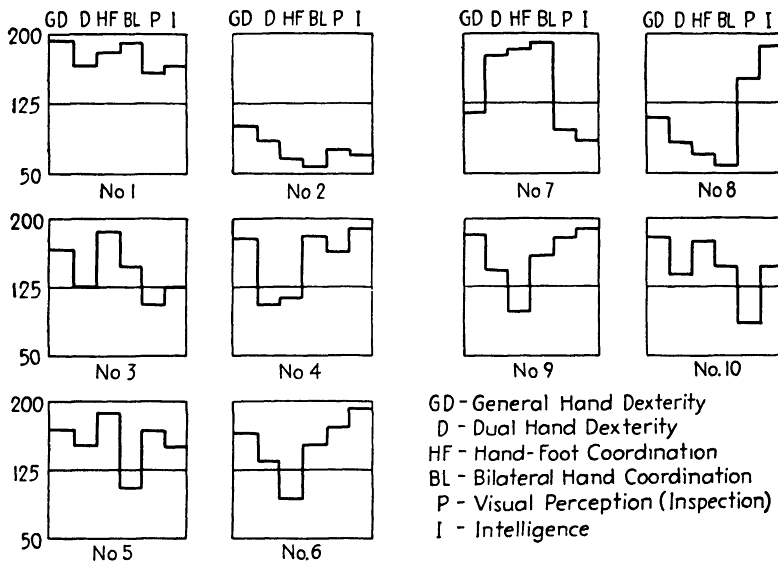

FIG. 30.—Profiles of industrial abilities, ten employees.

abilities that should be taken account of in assignment, to make certain that they are not placed on jobs for which they are specifically unfitted.

CHOOSING JOBS FOR TEST DESIGN

The number of tests to be designed is seen to depend upon the number of types of jobs for which operators are to be selected. The beginning is usually made with the jobs that are most difficult to fill with suitable employees or for which training is most difficult, prolonged, and expensive. In our own practice, we prefer to attack the most difficult job first. When we have solved the problems such a job presents, everyone is pretty well

convinced that the method will be applicable with equal or better success to the many easier and simpler jobs.

The person attempting to apply this technique for the first time might be well advised to try a simple job first, particularly a short-cycle dual job or a small-item packing job of the bilateral variety. Inspection jobs should come last, especially those involving span of perception, auditory perception, etc. The development of suitable tests in this latter area requires a background in experimental psychology, physiology, and psychophysics that is possessed by very few of those who may want to use this technique.

No matter how difficult a job may be, or how difficult is the task of designing a suitable test for it, some immediate success in better selection, better performance, and reduced training time should result from a conscientious effort to apply the principles we have stated. Many refinements may have to be made in the test as first designed. This is almost inevitable, and should not be a discouraging factor.

The designing of such tests is an art in itself. Eventually this art, like all others, will have a body of principles to guide its practice. These principles originate in hypotheses regarding human abilities and belong originally in the realm of philosophy. When the hypotheses have been subjected to experimentation and have been found to hold—a procedure that belongs to the science of industrial aptitude testing—they can be expressed in principles. These principles will, in turn, serve as guides in practice, *i.e.*, in applying the art of selection by specially designed tests.

CHAPTER XIV

SUMMARY AND FORECAST

This exposition of the new technique of industrial testing is not complete. New discoveries are constantly emerging as we prosecute this line of experimentation. Progress is continuous both in the refinement of the existing techniques and in the development of the new methods required as we attack fresh problems.

SUMMARY

It may be well to summarize what we believe to be the outstanding achievements resulting from our efforts thus far:

1. *Tests Designed from Time- and Motion-study Data.*—This is the outstanding and original contribution of this approach to industrial aptitude testing. The application of this method gives positive assurance that a properly designed test will do what it is supposed to do, *i.e.*, that it will select those who have in highest degree the innate capacities for the job, or related jobs, from which the design data are derived. To reiterate, the question "Will it work?" need no longer be asked. We know it will work. We need now only ask, "How well will it work?"

2. *The Discovery of New Areas of Human Individual Differences Significant for Success on Industrial Jobs.*—The list of these now numbers more than a dozen items, with every reason to believe that it may be extended to several dozen more. Outstanding on the list are dual coordination, bilateral coordination, hand-foot coordination, and the various perceptual areas. Many others of importance will continue to emerge.

3. *Percentage Efficiency Scoring for Tests.*—By the new technique of converting raw scores into percentage efficiency scores, it is possible to make a forecast of the eventual performance of the testee on the job for which the tests are designed. This

means performance after training and with suitable job setup, supervision, and incentives. It is possible by this procedure to establish an objective toward which the accomplishment of the individual on the job may be directed.

4. *The Test as Criterion.*—Since percentage efficiency scores may be used as a basis for predicting success on the job, the same measure will serve as a criterion against which to measure the success of supervision, and of management generally, in facilitating the indicated achievement. The basic argument for this procedure is that the test affords the most valid criterion, because of its inherent relationship with the job itself, and at the same time the most reliable and stable measurement of the ability called for.

5. *Separation of Motor and Perceptual Measurements.*—The deliberate separation of these two types of measurements has resulted in such notable success in the selection of inspectors and of operator-inspectors that discussion seems superfluous. This was such an obvious procedure in industrial testing that its belated introduction as a technique is hard to understand.

6. *The Motor-driven Test of Differential and Variable Speeds.*—This is a new technique that has already met with marked success in the selection of machine operators. It points the way in design for an infinite number of similar tests for operators of machine tools, machines, and vehicles. The most difficult feature of its development is to insure that it is measuring innate abilities rather than acquired skills. The basic principle in such a test can and will be extended in the several areas of perception, and in both perceptual and motor rhythm.

7. *Nonverbal Perception Tests.*—Visual Perception Tests A and B have demonstrated, by the unique success that has attended their application, the importance of a vast new area of nonverbal perception. This technique is the most logical one for the selection of salesmen, executives, engineers, and many others. The line of approach is now quite clear in devising highly effective tests for the selection of inspectors of cloth, tin plate, glass, bank-note paper, and other products. These developments await only the effective demand of the producers of such goods.

8. *Percentage Appraisal of Success, rather than Correlation.*—Tests designed by the method herein described are intended

to select the best and to eliminate the worst applicants in terms of the special abilities measured. We are primarily concerned with the proportion of the selected group who make good on the job. We are not concerned with those rejected for this job, but who may be, as we have demonstrated, quite well suited for jobs requiring other abilities. We have contended that the coefficient of correlation is a cumbersome and misleading generalization, and that it should be abandoned in favor of a simple percentage figure. It is our observation that failures among the persons whose measurements on the tests indicate high potential success on the job are almost always due to personality factors which lie outside the scope of the test and which were not identified in the personal interview procedures. In other words, the failures are due primarily to temperament and other aspects of personality, rather than to lack of the basic work abilities, the presence of which has been identified and guaranteed by the test results.

9. *The Measurement of Accident-proneness.*—This technique requires further investigation because of the probability that it may be easily and economically used to achieve substantial reductions in the number of accidents experienced among the work force. Such reductions seem to be largely independent of any amount of safety training. They can be achieved even after an intensive safety program has been in progress for years.

10. *Tests vs. Leveling.*—The leveling procedure in time-study technique, because of its subjective nature and lack of objective measurements of the skill and effort factors separately, is one of the outstanding limiting factors in the reliability of time studies. We contend that test results expressed in terms of percentage efficiency give a measure, under conditions of maximum incentive, of the basic innate abilities required, which can be readily used in place of the subjective and unreliable estimates. It is now up to the specialists in time and motion study to develop a method whereby these highly valid and reliable test scores, under conditions of maximum incentives, can be converted into usable terms at the optimum conditions prevailing on the job.

Forecast

The foregoing summarization foreshadows what can be expected of this technique in the immediate future. Even as this is written, new facts are being added to the body of experience from which the present procedures have been elaborated. No more propitious time than the present has ever existed for extending the experimentation along the lines of this technique. In the military establishment alone, with its acute need for effective selection and rapid training of the multitudes of men who are required to handle the new mechanized equipment, there is justification for setting up an extensive research laboratory in close proximity to the schools that are experimenting with the new weapons and equipment and for testing units in the field at every point where training is in progress. Why this has not been done already in view of the notable success that has attended the use of this technique in industry is difficult to explain. That it will be done eventually—it is hoped not too belatedly—is almost a certainty.

There is every reason to believe that when leaders in the labor movement become aware of the increased earnings and reduced effort that result from the application of this technique, they will not only urge, but actually demand, that it be used in the industries with which they are concerned. From a strictly nonpartisan, scientific viewpoint, it is believed that proper application of this technique has at least as much to offer labor as it has to offer management.

The further extension of the present developments may be expected to occur within two somewhat different areas. Experience thus far has shown that it is possible to design a standard battery of three dexterity tests and two perceptual tests that will cover a large proportion of jobs in industry. These tests will be far better than any combination of pencil-and-paper tests and other tests of a more pronounced psychological character. The battery of performance tests, particularly the dexterity tests, will be physiological rather than psychological. The results will be expressed in terms of anatomical and physiological capacities and disabilities, affected only in minor degree

by the greater variability of mental processes of such categories as reasoning, imagination, and intelligence.

The other line of development will include the construction of special tests where the number of jobs involved and the nature of the work justify this initial effort and expense. Such tests, designed from the results of time and motion studies and the other techniques indicated herein, will be better instruments of selection than any of the more general standard job tests. They will also be far superior to any selected battery of pencil-and-paper tests.

This technique has now arrived at the point where the whole operation of analysis of a job for test elements, design and construction of the test, and its standardization on a sample group requires only a few days. Testing time per employee is limited to not more than one hour, and often to a shorter period. If the test is properly designed on the basis of accurate data, if it is properly applied, and if the results are correctly interpreted, it must work. It will yield results that are both reliable and valid—results that may be used with confidence in the selection of employees.

It is hoped that all those who are in a position to experiment with tests of this character will undertake their development and use. If the principles are correctly followed, there should be few errors, and these not of a serious nature. Speed in the development of such tests and insight into their design and use will come with experience. Such successful experience will more than justify the time and effort spent in reading this book.

INDEX

A

Ability(ies), human, 10, 20*ff.*, 91-92
average, 4, 21
basic, testing of, 8-9, 22-23
innate, measurement of, 14
motor, range of, 8-9
new areas of, 25-26, 122
profiles of, 130-131
and reassignment of operators, 92
and shorter day, 91
Accident hazard, classifying jobs in terms of, 122
Accident index, 115
comparison of, with test scores, 117-119
computation of, 116
relation of, to earnings, 120-121
selection by test, a factor in reduction of, 120-122
(*See also* Selection, by test)
Accident-proneness, 10, 54, 98-99
Choice Reaction Time Test, relation to, 114
Dotting Test, relation to, 114
effect of safety education and training on, 113, 122
hypothesis, 54, 119-120
implications of, 122
individual differences in, 113-114
intelligence, relation to, 114
measurement of, 113*ff.*, 135
Pursuit Meter Test, relation to, 114
Accidents, control of, through safety training, 113
fear, relation to, 98-99
severity of, 116-117
tests, relation to, 114-119
American Machinist, 115*n.*
Aptitude(s), engineering, 52-53
high, effects of, 102
low, effects of, 96, 102
mechanical, tests of, 5
and skill, 102*ff.*
Assembly, simultaneous, 60
Average, the, as ideal, 15, 97

B

Bilateral Hand Test, 82-83
correlation of scores on, with other tests, 82
operation of, 82
Bingham, W. V., findings of, 113, 113*n.*
Brain, structure of, 100-101
motor area, 100
supermotor area, 100
visual perceptual areas, 100
relation to visual sensory areas, 35, 100
visual sensory areas, 100
British Industrial Health Research Board (formerly British Industrial Fatigue Research Board), 113*n.*, 114

C

"Card sense," 51
Case Inspection Test, 41
error scores on, 43
relation of, to dexterity tests, 27, 42
operation of, 40-41, 116
retest on, 44
time scores on, 43
relation of, to dexterity tests, 27, 42
Character and Personality, 114*n.*

Choice Reaction Time Test, 114
Controlled Turning Test, 61, 62
  operation of, 62, 63
  reliability of, 64
  savings on, 63, 64, 65
Coordination, bilateral hand, area of individual differences, 16, 83
  dual, 60*ff.*
    individual differences in, 60, 66
    tests of, 61-66
      (*See also* Dual operation)
  dual hand, 12
    area of individual differences, 16
    category of test design, 13
  dual hand-and-one-foot, area of individual differences, 16
  general body, area of individual differences, 16
    category of test design, 13
  hand-foot, 10, 14, 24, 27*ff.*, 78
    category of test design, 13
    implications of, for industrial engineering, 34
    individual differences in, 27, 66
    relation of, to accidents, 32, 33
  motor, measurement of, 5, 6
Correlation, coefficient of, 11-12, 125
  inter-, of aptitude test scores, 27
  percentage appraisal as opposed to, 134-135
  probable error, relation to, 12
Critical score, the, 44, 126-127
  influence of labor market on, 85, 126

D

Design of tests, categories of jobs for purposes of, 13
  choosing jobs for, 131-132
  considerations in, 129-130
  elements in, 13
  principles of, 16-18, 132
  training for, 18-19
Detroit Mechanical Aptitude Test, 73
  correlation of scores on, with Machine-operator Test, 70
Detroit Mechanical Aptitude Test, scores on, relation to foreman's ratings, 72
Dexterity, design of tests of, 129-130
  general finger, area of individual differences, 16
  general hand, 27
    category of test design, 13
Dotting Test, 114
Drake, Charles A., investigations of, 114*n.*, 115*n.*
Dual Hand-and-foot Test, 30, 31
  intercorrelation of scores on, with other measures, 29
  operation of, 29
Dual Hand Test, 81-82
  operation of, 81-82
Dual operation, 10, 14, 24, 94
  ability, specialized function, 60-61
  mechanical aids in, 66
  savings by, 60, 63-64, 66, 81
    percentage, computation of, 61-62

E

Eagle Pencil Company, experimentation at, 12-13, 27-28, 29, 39-44
Earnings, possible criterion for measuring effectiveness of test, 11, 14
Effort ratings, 4
Engineering schools, responsibility of, 3
Eugenical News, 113*n.*

F

Factory Management and Maintenance, 36*n.*, 115*n.*
Farmer, E., and Chambers, E. G., investigations of, 114, 120, 121
Fatigue, decrease in, 1-2
  repetitious work and, 90-91
Fear, of work, 98
  and accidents, 98-99
Fixtures, design of, 24, 32-33
  strength of operator, a factor in, 33
  (*See also* Machine design)

Foreman's rankings, 116
"halo" effect in, 13
possible criterion for measuring effectiveness of test, 11, 13-14
relation of, to scores on Pin Board E, 8-9
to scores on other tests, 70-72
reliability of, 14
validity of, 14
Foremen, selection of, 92-93, 95

G

Gilbreths, work of, 1, 2, 12
Greenwood, M., and Wood, H. M., investigations of, 113, 113*n*.

H

Hand dexterity tests, 78-80
Hand-foot Test I, 28
intercorrelation of scores on, with other measures, 27-28
operation of, 28
range of scores on, 29
Hand-foot Test II, 77-78
"Hand" of fabrics, tests for, 57-58
Henmon-Nelson Test of Mental Ability, 45
Human factor, measurement of, 1, 4, 10
probable error in handling of, 7
weakness of management in dealing with, 1

I

Indexing, on job and on test, 127, 129
Individual differences, human, in abilities and aptitudes, 3
areas of, 13, 16, 24, 133
bearing of, on performance, 13, 23
new areas of, 16, 25-26, 133
methods engineering, relation to, 15
Industrial Fatigue Research Board Reports, 113*n*., 114*n*.
Inspection, 10, 35*ff*., 49, 51
auditory, 35
gustatory, 35
kinesthetic, 80
olfactory, 35
and quality control, 47
tactual, 35
tests, 35, 123-125
(*See also* Tests, perception)
visual, 35
(*See also* Perception)
Inspection ability, 14, 25
auditory acuity, relation to, 35
intelligence, relation to, 45
visual acuity, relation to, 35-36, 46
(*See also* Perceptual ability)
Inspection Test, 80-81
operation of, 80
reliability of, 80-81
Intelligence, general, 51
area of individual differences, 23
individuals of average, 21
necessity for effecting an organization around, 21
individuals of inferior, 22
individuals of superior, 21-22
social intelligence, relation to, 50-51
testing, history of, 4-5

J

Job analyses, basis for test design and specifications, 60, 73
Job testing, backgrounds of, 10, 11*ff*.
(*See also* Tests, job)
Johnson and Johnson, experimentation at, 29-30, 98
Journal of Engineering Education, 3*n*.

K

Kinesthetic Perception Test, 56, 57
operation of, 55-56
reliability of, 57

L

Learning ability, 10
  effects of low aptitude on, 96, 102
  influence of fear on, 98-99
  influence of previous failure on, 99
  other inhibitions to, 99-100
  tests and, 94*ff.*
Learning curves, 111-112
  combined scores, 110
  hypothesis, 106
  individual, 108, 109, 110
  for individual of high aptitude, 102
    of low aptitude, 102
Left-right Turning Test, 61, 63
  operation of, 62-63
  reliability of, 64
  savings on, 63, 64
Limen, 124
Link, Henry C., 76

M

Machine, as test, 30-31
Machine design, 17-18, 32-33
  limitations of human structure on, 23-24
Machine-operator Test, 67*ff.*
  intercorrelation of scores on, with other measures, 70-71, 72
  operation of, 68-69
  ranks on, relation to foreman's ratings, 71, 72
  reliability of, 70, 73
Management, as practical art, 1*ff.*
  scientific, status of, 1-2
Methods engineering, relation of, to management, 15
  as science and art, 15-16
    (*See also* Selection)
"Minimal cues," 49-50
  individual differences in response to, 57
  measurement of ability to react to, 57
Mogensen, Allan H., 12
Monotony, 90, 91
Motor reactions, effect of alcohol on, 54, 120
  effect of distractions on, 120
  effect of fatigue on, 54, 120
  effect of increase of speed of, 53
  individual differences in, 53
Motor Rhythm Test, 76-77
  operation of, 76
  reliability of, 76-77
  scores on, correlation with other measures, 27, 76, 77
    distribution of, 87

N

National Safety Council, 99
Newbold, E. M., investigations of, 113, 113*n.*

O

O'Connor Tweezer Dexterity Test, operation of, 115
Oleen, Holger D., 36*n.*, 124
Operation cycle, 2, 16, 17, 31
  analysis of, as basis for design of tests, 12, 83
  length of, effect of, on percentage efficiency, 87-88
  long, 87-88, 89, 90
    forecasting production on, 129
  serialization of elements of, 100
  short, effect of, on savings by dual operation, 60
    forecasting production on, 129
    implications of, for production, 89
Otis Self-administering Test of Mental Ability, 45

P

Percentage efficiency, a possible criterion for measuring effectiveness of test, 14
Perception(s), 10, 25, 48*ff.*
  accuracy of, range of ability in, 6
  application of facts of, 49-50

Perception(s), auditory, 12, 35, 46
dimensions of, 46
and quality control, 54
test of, 46-47
auditory acuity, relation to, 35
definition of, 6
effect of alcohol on, 54, 120
effect of distractions on, 120
effect of drugs on, 54
effect of emotional states on, 54
effect of fatigue on, 54, 120
engineering aptitude and, 52-53
individual differences in, 53, 57, 125
intuition, relation to, 50
kinesthetic, 12, 55, 56-58, 71
and quality control, 54
measurement of, 5, 49
nonverbal, 52
reasoning, relation to, 49
sales, tests of, 50-51
sensation, relation to, 49
sensory acuity, relation to, 35
social intelligence, aspect of, 50-51
of space and time relationships, 70
speed of, range of ability in, 6
tactual, 12, 54-55, 58
in three dimensions, 38, 119
in two dimensions, 38, 119
visual, 58, 71, 100
tests of (*see* Visual perception tests)
visual acuity, relation to, 6, 35, 36, 46, 48-49, 120
Perceptual ability, area of individual differences, 16
background, relation to, 45
category of test design, 13
illumination, relation to, 45, 50
sense organs, relation to, 25
visual, kinds of, 36
discrimination of small differences, 36
tests for, 36-44, 123
span of perception, 36
test for, 44-45
deterioration of, 46
Perceptual ability, effect of color-weakness and color-blindness on, 49
periodic checkups of, 46
(*See also* Visual perception tests)
Perceptual measurements, separation of, from motor measurements, 134
Performance, level of, 94-95
optimum, 24-25
Personnel, 114*n*.
"Physiological limit," in industry, 94, 111
Pin Board E, 61
operation of, 62-63, 115
relation of scores on, to foreman's rankings, 8-9
reliability of, 64
savings on, 63, 64
Pin Board I, 78, 79
operation of, 78-80
Pin Board II, 79
operation of, 80
"Plateaus," learning, 99-100
Positioning, pre-, 127
on test and on job, 16, 31, 127, 129
Probable error, 7, 9, 10
as basis for forecasts, 7
coefficient of correlation, relation to, 12
concept of, 7-8
Production, a possible criterion for measuring effectiveness of test, 11, 14
Psychographs, 117
Pursuit Meter Test, 114

Q

Quality, a possible criterion for measuring effectiveness of test, 14
Quality control, bearing of perception on, 54
inspection and, 47

R

Rating scales, weaknesses of, 4
Reduced cues, 49-50
  individual differences in response to, 57
  measurement of ability to react to, 57
Rhythm, area of individual differences, 16
  of work, individual differences in, 24-25
Right-right Turning Test, operation of, 115
  scores on, relation to accident index, 118-119

S

Scientific Monthly, The, 114*n.*
Scores, critical, 44, 126-127
  percentage efficiency, 86-87, 89, 133-134
    advantages of, 86-87
    relation to length of cycles, 87-88
  percentile, 85-86
    objection to use of. 86
  standard, 86
Selection, effect of labor market on, 85, 126
  of foremen, 92-93, 95
  of inspectors, 4
    relation of intelligence to, 45
  industrial employee, requirements of, 12
  problem of, 96
  relation of earnings to, 111-112
  of salesmen, 51-52
  as science and art, 15-16
  by test, 11, 51, 97
    as a factor in accident reduction, 98-99, 114-115
  use of industrial ability profile in, 130-131
Self-expression, in industry, 20
Sensation, as distinguished from perception, 6
Sewing-machine Test, 75
  operation of, 76
Skill(s) criteria of, in performance, 111
  effects of low aptitude on, 96, 102
  inhibitions to, 99-100
  motor, and production, 85*ff.*
  muscular, experiment in learning of, 102*ff.*
  new, development of, 100-101
  ratings, 4
  transfer of, 101
Spiral Inspection Test, 41
  error scores on, 42
    relation to dexterity tests, 27
  operation of, 39-40, 115-116
  relation of scores on, to accident index, 118, 119
  time scores on, 42
    relation to dexterity tests, 27

T

Tactual Perception Test, 55
  operation of, 54
Taylor, Frederick W., 1
Test(s), aptitude, design of, 18
    development of, 9
    intercorrelations of scores on, 27
    legal, 49
    mechanical, 5, 70-71
    sales, 49, 51-52
  battery of, 11, 12, 136
  Bilateral Hand, 82-83
    correlation of scores on, with other tests, 82
    operation of, 82
  Case Inspection, 41
    error scores on, 43
      relation of, to dexterity tests, 27, 42
    operation of, 40-41, 116
    retest on, 44
    time scores on, 43
      relation of, to dexterity tests, 27, 42
  construction of, 123*ff.*
  Controlled Turning, 61, 62

Test(s), Controlled Turning, operation of, 62, 63
  reliability of, 64
  savings on, 63, 64, 65
 as criterion, results of, for measuring efficiency of productive process, 14-15, 85, 134
 design of, art of, 17, 18, 132
  categories of jobs for purposes of, 13
  choosing jobs for, 131-132
  considerations in, 129-130
  dexterity, 129-130
  for discrimination of small differences, 36-44
  elements in, 13
  principles of, 16-18, 132
  a simple dual, 127, 129
  for span of perception, 44, 45
  training for, 18-19
 Detroit Mechanical Aptitude, 73
  correlation of scores on, with Machine-operator Test, 70
  relation of scores on, to foreman's ratings, 72
 Dual Hand, 81-82
  operation of, 81-82
 Dual Hand-and-foot, 30, 31
  intercorrelation of scores on, with other measures, 29
  operation of, 29
 hand dexterity, 78-80
 Hand-foot I, 28
  intercorrelation of scores on, with other measures, 27-28
  operation of, 28
  range of scores on, 29
 Hand-foot II, 77-78
 inspection, 35, 123-125
  (*See also* Tests, perception)
 Inspection, 80-81
  operation of, 80
  reliability of, 80-81
 intelligence, relationship between results on, and performances on jobs, 4
  in selection of inspectors, 4
  in selection of management personnel, 4-5
Test(s), intelligence, social, 49
 job, 6, 75*ff.*
  (*See also* Job testing)
 Kinesthetic Perception, 56, 57
  operation of, 55-56
  reliability of, 57
 and learning ability, 94*ff.*
 Left-right Turning, 61, 63
  operation of, 62-63
  reliability of, 64
  savings on, 63, 64
 Machine-operator, 67*ff.*
  intercorrelation of scores on, with other measures, 70-71, 72
  operation of, 68-69
  relation of ranks on, to foreman's ratings, 71, 72
  reliability of, 70, 73
 motor, 17, 44
  relation of scores on, to accident index, 117
 motor-driven, 57, 67, 134
 Motor Rhythm, 76-77
  correlation of scores on, with other measures, 27, 76, 77
  distribution of scores on, 87
  operation of, 76
  reliability of, 76-77
 O'Connor Tweezer Dexterity, operation of, 115
 pencil-and-paper, 11
  utility of, in selecting operating personnel, 22, 23
 perception, 134
  for discrimination of small differences, 36-44
   Original Visual Perception Test, 124
   Visual Perception Test A, 37, 76
    intercorrelation of scores on, with Machine-operator Test, 70, 71, 73
    scores on, relation to foreman's ratings, 72
    tendency to bimodality of scores on, 53

Test(s), perception, Visual Perception Test B, 37, 38, 76
 scores on, intercorrelation of, with Machine-operator Test, 70, 71, 73
 relation of, to foreman's ratings, 72
 tendency to bimodality of scores on, 53
 Visual Perception Test C, 37, 39
 Visual Perception Test D, 37, 40
 scores on, relation to accident index, 117
 for span of perception, 44, 45
performance, 10
 construction of, 123*ff*.
 in measurement of ability to work rhythmically, 24
 in selection of operating personnel, 23
Pin Board E, 8, 61
 operation of, 62-63, 115
 reliability of, 64
 savings, 63, 64
 scores on, related to foreman's ranking, 8, 9
Pin Board I, 78, 79
 operation of, 78-80
Pin Board II, 79
 operation of, 80
possible criteria for measuring effectiveness of, 13-15
 earnings, 11, 14
 foreman's rankings, 11, 13, 14
 percentage efficiency, 14
 production, 11, 14
 quality, 14
reliability of, 44, 84, 125
Right-right Turning, operation of, 115
 scores on, relation to accident index, 118, 119
sensory, 48
Sewing-machine, 75
 operation of, 76
Spiral Inspection, 41
Test(s), Spiral Inspection, error scores on, 42
 relation of, to dexterity tests, 27
 operation of, 39-40, 115-116
 scores on, relation to accident index, 118-119
 time scores on, 42
 relation to dexterity tests, 27
standardized, 11, 13
Tactual Perception, 55
 operation of, 54
Textile Thread, 128
 operation of, 129
trade knowledge, 5
validity of, 84, 125
verbal, 4-5
 ability identified by, 5
work-limit type, 17, 115
 handling of scores on, 85, 86
Textile Thread Test, 128
 operation of, 129
Textile World, 57*n*.
Therblig Analysis, 12-13
 basic elements for test design, 12-13
Therbligs, on test and on job, 16
Time and motion study, 2-4
 attitude toward, 2-3
 data, as basis for test design, 133
 leveling procedure in, 135
 trial and error in, 3
Time-study engineers, 2, 3, 4
 and aptitude testing, 3*n*.
 responsibility of, 3, 9
Training, foremanship, 95
 safety, 113
 (*See also* Tests, and learning ability)

V

Visual perception tests, 36-44, 45
 Original Visual Perception Test, 124
 Visual Perception Test A, 37, 76
 intercorrelation of scores on, with Machine-operator Test, 70, 71, 73

Visual perception tests
  Visual Perception Test A, relation of scores on, to foreman's ratings, 72
    tendency to bimodality of scores on, 53
  Visual Perception Test B, 37, 38, 76
    intercorrelation of scores on, with Machine-operator Test, 70, 71, 73
    scores on, relation to foreman's ratings, 72
    tendency to bimodality of scores on, 53
Visual perception tests, Visual Perception Test C, 37, 39
  Visual Perception Test D, 37, 40
Vocational guidance counselors, plaint of, 20, 21

W

Wonderlic's adaptation of Otis Test, scores on, intercorrelation of, with Machine-operator Test, 70, 71-72
    relation of, to foreman's ratings, 72
Work Simplification program, 12